Journal
of
The Franklin Institute
Devoted to Science and the Mechanic Arts

Vol. 267 MAY, 1959 No. 5

THEORY OF THE A. F. JOFFE METHOD FOR RAPID MEASUREMENT OF THE THERMAL CONDUCTIVITY OF SOLIDS*

BY

W. F. G. SWANN[1]

ABSTRACT

The figure below represents a block B of high thermal conductivity fastened to a specimen S of length l parallel to the axis of x, and of lower conductivity k. At $t = 0$, all is at constant temperature. Then, suddenly, the right hand side of S is brought to a temperature τ_∞ and kept there. If θ is the excess of the temperature at some point over τ_∞, the general problem consists in finding θ as a function of x and t.

The simple theory assumes that θ, at $x = 0$, decreases so slowly with time that the temperature gradient in S is, at each instant, independent of x and equal to θ/l. This assumption leads to $\theta = \theta_0 e^{-\lambda_0 t}$, where θ_0 is the initial value of θ, and $\lambda_0 = k/Msl$, M and s being, respectively, the total mass per unit area of B projected on the plane of contact, and the specific heat of block B.

The more exact theory gives

$$\theta = \sum_0^\infty A_s e^{-\lambda_s t}\left[\cos\frac{\psi_s x}{l} - G_s\psi_s \sin\psi_s\frac{x}{l}\right]$$

where the constants A_s are to be determined to give $\theta = \theta_0$ at $t = 0$. λ_s is given by $R\lambda_s/\lambda_0 = \psi_s^2$ where R is the ratio of the thermal capacity of the specimen to that of B, ψ_s is a solution of the equation $G_s\psi_s \tan\psi_s = 1$, and G_s is a constant function of s whose form depends upon the nature of the boundary conditions at $x = 0$—the place of junction of B and S. For the case where there is perfect thermal contact at $x = 0$, $G_s = 1/R$ and is independent of s. For the case where there is thermal resistance at $x = 0$, $G_s = [R + h\psi_s^2/hl]^{-1}$ where h is defined by

$h\Delta\theta$ = heat transferred across the boundary per cm^2 per sec.

and $\Delta\theta$ is the temperature drop. It turns out that, at the junction, θ is of the form

$$\theta = A_0 e^{-\lambda_0 t} + A_1 e^{-\lambda_1 t} + A_2 e^{-\lambda_2 t} + \cdots,$$

where λ_1, λ_2, λ_3, etc. are increasingly larger than λ_0, and by very substantial amounts,

* Since this paper was completed, there came to the author's attention an unpublished copy (a preprint) of a paper by M. A. Kaganov dealing with the theory of Joffe's method.

[1] Director, Bartol Research Foundation of The Franklin Institute, Swarthmore, Pa., and a Senior Staff Advisor for The Franklin Institute Laboratories for Research and Development.

(Note—The Franklin Institute is not responsible for the statements and opinions advanced by contributors in the JOURNAL.)

so that, except near $t = 0$, the term involving λ_0 is all important, and $\theta = A_0 e^{-\lambda_0 t}$ at $x = 0$. It then results that, for R small, and k reasonably large

$$\frac{\lambda_0}{\lambda_a} = \frac{1}{1 + \frac{R}{3}} \quad \text{for thermal junction resistance absent}$$

and

$$\frac{\lambda_0}{\lambda_a} = \frac{1}{1 - \frac{k}{hl} + \frac{R}{3}} \quad \text{for thermal resistance included.}$$

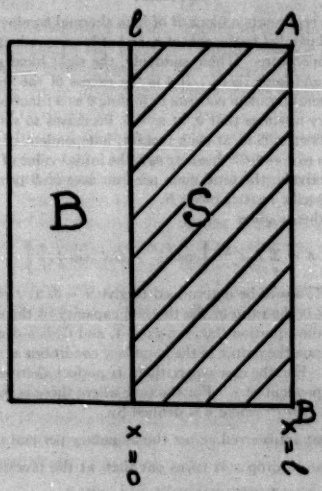
λ_0 is the experimentally measured value, so that we should have, for the two cases, respectively,

$$\frac{k}{Msl} = \left(1 + \frac{R}{3}\right)\lambda_0$$

$$\frac{k}{Msl} = \left(1 - \frac{k}{hl} + \frac{R}{3}\right)\lambda_0.$$

The elementary theory omits the quantities in parentheses.

The actual determination of the A_s's to complete the expression for θ as a function of x and t is carried out for the case where thermal resistance at the junction is neglected.



INTRODUCTION

A method used extensively for measuring the thermal conductivity of solids, including semi-conductors, in connection with thermoelectric research, is dependent upon the following principle:[2]

[2] A. F. JOFFE, *Can. J. Phys.*, Vol. 34, p. 1342 (1956).

A block B of high thermal conductivity is fastened to a slab S represented by the shaded portion of the figure. The thermal conductivity of S is low compared with that of the mass B, so that we may assume, for B, a temperature which, though time dependent, is constant throughout the mass although the temperature may vary throughout S in a manner determined by the nature of the problem.

Our problem is such that, at $t = 0$, everything is at the constant temperature τ_0. Then, suddenly, (at $t = 0$) the face AB of S is brought to the temperature τ_∞, in actuality the temperature of melting ice, and kept there for all time. We desire to know the temperature τ of S at each subsequent instant, as a function of x, measured from left to right, the plane of junction with B being at $x = 0$. We assume the dimensions perpendicular to the axis of x to be so large that we may treat the problem as one-dimensional. Moreover, we assume that B loses heat only at the plane of junction with S.

We make the following definitions:

M = mass of B divided by the area of its junction with S,
s = specific heat of the material of the mass B,
l = thickness of slab S,
k = thermal conductivity of the material of the slab S,
σ = specific heat of the material of the slab S,
α^2 = ratio of thermal conductivity to product of specific heat and density for S,
$\Delta\theta$ = drop in temperature at the junction of S and B on account of a finite thermal resistance at that junction, and
h = coefficient of $\Delta\theta$ in the expression for the heat conducted per cm^2 per second across the junction.

The solution of this problem obviously provides a method for determining the thermal conductivity of S in terms of the temperature-time relationship for B, and of the relevant dimensions and other parameters of the system.

The basic differential equation controlling the temperature τ in the slab S is

$$\frac{\partial \tau}{\partial t} = \alpha^2 \nabla^2 \tau. \tag{1}$$

It will be convenient to define a quantity θ by

$$\theta = \tau - \tau_\infty$$

so that θ is the excess of τ over τ_∞. In terms of θ, Eq. 1 becomes

$$\frac{\partial \theta}{\partial t} = \alpha^2 \nabla^2 \theta. \tag{2}$$

Conditions Governing the Solution

We have:

A. θ must satisfy (2) for all values of t over the range $0 < x < l$
B. $\theta = \theta_0$ at $t = 0$ for all values of x between $x = 0$ and $x = l$
C. At $x = l$, $\theta = 0$ for all time.
D. At $t = \infty$, $\theta = 0$ for all values of x, from $x = 0$ to $x = l$
E. At $x = 0$

$$Ms\frac{\partial\theta}{\partial t} = k\frac{\partial\theta}{\partial x} \tag{3}$$

or

$$Ms\frac{\partial\theta}{\partial t} - \frac{Msk}{h}\frac{\partial^2\theta}{\partial t\partial x} = k\frac{\partial\theta}{\partial x} \tag{3.1}$$

according as we do not or do take account of a thermal resistance between the junction of S and B.[3]

Customarily Used Approximate Solution

If we assume that Ms is very large, it is clear that, unless k is very large, θ, at $x = 0$, will diminish very slowly with the time. If we assume that, at each instant, $\partial\theta/\partial x$ assumes, in S, a value independent of x and equal to $-\theta/l$, Eq. 3, neglecting thermal resistance at the junction, yields[4]

$$Ms\frac{\partial\theta}{\partial t} = -\frac{k\theta}{l}$$

$$\theta = \theta_0 e^{-kt/Msl} \tag{4}$$

which expression enables k/Ms to be determined through knowledge of the experimentally measured θ,t relationship.

---

[3] In terms of the notation in the system of definitions, we have

$$Ms\frac{\partial(\theta + \Delta\theta)}{\partial t} = k\frac{\partial\theta}{\partial x} = -h\Delta\theta$$

where θ is the temperature at $x = 0$, so that the temperature of the surface of B at $x = 0$ is $\theta + \Delta\theta$. From these relations we deduce immediately

$$Ms\frac{\partial\theta}{\partial t} - \frac{Msk}{h}\frac{\partial^2\theta}{\partial l\partial x} = k\frac{\partial\theta}{\partial x}.$$

[4] It is clear that $\partial\theta/\partial x$ must, in actuality, always differ from the constant value $-\theta/l$ to the extent necessary to provide for the fact that however slowly θ diminishes with time at $x = 0$, heat must flow out of the various parts of S at a corresponding rate sufficient to insure the continued fall in the average temperature of S. All of the heat flowing out of S *on this account* passes through $x = l$, while none of it passes through at $x = 0$. It is manifest therefore that $d\theta/dx$ at $x = l$ must differ from that at $x = 0$ to the extent necessary to harmonize matters.

THE GENERAL THEORY

In view of the uncertainty as to the magnitudes of the errors involved in the assumptions leading to (4), and of the general applicability of the method, it seems worth while to develop the theory of the experiment in a general manner. We shall, however, retain the assumption that the conductivity of the material of the mass B is so good that the whole of the mass may be assumed to be of uniform temperature at each instant.

DEVELOPMENT OF THE SOLUTION

Convenient Relations

It will be convenient to note the following relations between the constants of the system.

We define λ_a as

$$\lambda_a \equiv \frac{k}{Msl}. \tag{5}$$

Thus λ_a is the reciprocal of the "time constant" of the approximate solution, as exemplified by (4).

We define R as

$$R \equiv \frac{\text{heat required to raise } S \text{ one degree centigrade}}{\text{heat required to raise } B \text{ one degree centigrade}} \tag{6}$$

$$R = \frac{\rho\sigma l \text{ (area of contact)}}{Ms \text{ (area of contact)}} = \frac{\rho\sigma l}{Ms}. \tag{7}$$

Again, since $\alpha^2 = k/\rho\sigma$, we have

$$\frac{k}{Ms\alpha^2} = \frac{k\rho\sigma}{Msk} = \frac{R}{l}. \tag{8}$$

Also

$$\alpha^2 = \frac{k}{\rho\sigma} = \frac{k}{Msl}\cdot\frac{Msl}{\rho\sigma} = \frac{\lambda_a l^2}{R}. \tag{9}$$

Analytical Procedure

Let us write

$$\theta = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots \tag{10}$$

where the a's are functions of the time. Then (2) demands that

$$(\dot{a}_0 + \dot{a}_1x + \dot{a}_2x^2 + \dot{a}_3x^3 + \dot{a}_4x^4 + \cdots) = \alpha^2(2a_2 + 3\cdot 2a_3x + 4\cdot 3a_4x^2 + 5\cdot 4a_5x^3 + \cdots).$$

Since this must hold for all values of x, we require

$$\left.\begin{aligned} a_2 &= \frac{\dot{a}_0}{\lfloor 2\alpha^2}; & a_3 &= \frac{\dot{a}_1}{\lfloor 3\alpha^3} \\ a_4 &= \frac{\dot{a}_2}{4\cdot 3\alpha^2} = \frac{\ddot{a}_0}{\lfloor 4\alpha^4}; & a_5 &= \frac{\dot{a}_3}{5\cdot 4\alpha^2} = \frac{\ddot{a}_1}{\lfloor 5\alpha^4} \\ a_6 &= \frac{\dot{a}_4}{6\cdot 5\alpha^2} = \frac{\dddot{a}_0}{\lfloor 6\alpha^6}; & a_7 &= \frac{\dot{a}_5}{7\cdot 6\alpha^2} = \frac{\dddot{a}_1}{\lfloor 7\alpha^6} \\ &\cdots\cdots\cdots & &\cdots\cdots\cdots \\ a_{2s} &= \frac{1}{\alpha^{2s}\lfloor 2s}\frac{d^s a_0}{dt^s}; & a_{2s+1} &= \frac{1}{\alpha^{2s}\lfloor 2s+1}\frac{d^s a_1}{dt^s} \end{aligned}\right\}. \quad (11)$$

Thus, from (10)

$$\theta = a_0 + \frac{\dot{a}_0 x^2}{\lfloor 2\alpha^2} + \frac{\ddot{a}_0 x^4}{\lfloor 4\alpha^4} + \frac{\dddot{a}_0 x^6}{\lfloor 6\alpha^6} + \cdots + a_1 x + \frac{\dot{a}_1 x^3}{\lfloor 3\alpha^3} + \frac{\ddot{a}_1 x^5}{\lfloor 5\alpha^4} + \frac{\dddot{a}_1 x^7}{\lfloor 7\alpha^6} + \cdots. \quad (12)$$

Now condition C requires that

$$a_0 + \frac{\dot{a}_0}{\lfloor 2}\left(\frac{l}{\alpha}\right)^2 + \frac{\ddot{a}_0}{\lfloor 4}\left(\frac{l}{\alpha}\right)^4 + \frac{\dddot{a}_0}{\lfloor 6}\left(\frac{l}{\alpha}\right)^6 + \cdots + \alpha\left[a_1\left(\frac{l}{\alpha}\right) + \frac{\dot{a}_1}{\lfloor 3}\left(\frac{l}{\alpha}\right)^3 + \frac{\ddot{a}_1}{\lfloor 5}\left(\frac{l}{\alpha}\right)^5 + \frac{\dddot{a}_1}{\lfloor 7}\left(\frac{l}{\alpha}\right)^7 + \cdots\right] = 0. \quad (13)$$

We now try as a solution of this equation

$$a_0 = Ae^{-\lambda t}; \quad a_1 = -Be^{-\lambda t} \quad (14)$$

where A and B will become subsequently related to fit condition (3), or (3.1), whichever we choose to adopt.

$$A\left[1 - \frac{1}{\lfloor 2}\left(\frac{l\sqrt{\lambda}}{\alpha}\right)^2 + \frac{1}{\lfloor 4}\left(\frac{l\sqrt{\lambda}}{\alpha}\right)^4 - \frac{1}{\lfloor 6}\left(\frac{l\sqrt{\lambda}}{\alpha}\right)^6 + \cdots\right] - \frac{\alpha}{\sqrt{\lambda}}B\left[\left(\frac{l\sqrt{\lambda}}{\alpha}\right) - \frac{1}{\lfloor 3}\left(\frac{l\sqrt{\lambda}}{\alpha}\right)^3 + \frac{1}{\lfloor 5}\left(\frac{l\sqrt{\lambda}}{\alpha}\right)^5 - \cdots\right] = 0 \quad (15)$$

yielding

$$A\cos\left(\frac{l\sqrt{\lambda}}{\alpha}\right) - \frac{\alpha B}{\sqrt{\lambda}}\sin\left(\frac{l\sqrt{\lambda}}{\alpha}\right) = 0 \quad (16)$$

or, using (9)

$$A\left[\cos\left(\frac{R\lambda}{\lambda_s}\right)^{1/2} - \frac{Bl}{A}\left(\frac{\lambda_s}{R\lambda}\right)^{1/2}\sin\left(\frac{R\lambda}{\lambda_s}\right)^{1/2}\right] = 0. \tag{17}$$

Writing

$$\psi \equiv (R\lambda/\lambda_s)^{1/2} \quad \text{and} \quad \frac{Bl}{A}\left(\frac{\lambda_s}{R\lambda}\right) = G \tag{18}$$

$$\psi \tan\psi = 1/G. \tag{19}$$

B/A, and so G, become determined so as to fit condition (3), or (3.1), in accordance with our choice, and (19) then determines the possible values of ψ and so, through (18), the possible values of λ. G will, in general, involve λ. We shall return to these matters. For the moment we proceed to the general expression for θ as a function of x and t.

General Expression for θ as a Function of x *and* t

Writing λ_s for a typical value of λ, and A_s and B_s and G_s as the associated values of A and B and G, we have, from (14),

$$a_{0s} = A_s e^{-\lambda_s t}; \quad a_{1s} = -B_s e^{-\lambda_s t} = -\frac{A_s R\lambda_s}{l\lambda_s} G_s e^{-\lambda_s t} \tag{20}$$

as the associated possible values of a_0 and a_1 satisfying (13), and so condition C. Moreover, any values of a_0 and a_1 given by (20) satisfy (13). We can, therefore, write (12), our solution for θ, in the form

$$\theta = \sum_0^\infty A_s e^{-\lambda_s t}\left[1 - \frac{\lambda_s}{\lfloor 2}\left(\frac{x}{\alpha}\right)^2 + \frac{\lambda_s^2}{\lfloor 4}\left(\frac{x}{\alpha}\right)^4 - \frac{\lambda_s^3}{\lfloor 6}\left(\frac{x}{\alpha}\right)^6 + \cdots\right.$$

$$\left. - \frac{R\alpha\lambda_s^{1/2}}{l\lambda_s} G_s\left\{\lambda_s^{1/2}\left(\frac{x}{\alpha}\right) - \frac{\lambda_s^{3/2}}{\lfloor 3}\left(\frac{x}{\alpha}\right)^3 + \frac{\lambda_s^{5/2}}{\lfloor 5}\left(\frac{x}{\alpha}\right)^5 - \cdots\right\}\right]$$

$$\theta = \sum_0^\infty A_s e^{-\lambda_s t}\left[\cos\lambda_s^{1/2}\frac{x}{\alpha} - \frac{R\alpha\lambda_s^{1/2}}{l\lambda_s} G_s \sin\lambda_s^{1/2}\frac{x}{\alpha}\right].$$

Using (9) and (18)

$$\theta = \sum_0^\infty A_s e^{-\lambda_s t}\left[\cos\frac{\psi_s x}{l} - G_s\psi_s \sin\frac{\psi_s x}{l}\right] \tag{21}$$

where G_s is to be determined by the condition (3) or (3.1), as the case may be. This gives $\theta = 0$ at $t = \infty$ and so satisfies (D). It remains to

determine the A's so that they shall satisfy (B). Thus, we must have

$$\sum_0^\infty A_s \left[\cos \frac{\psi_s x}{l} - G_s \psi_s \sin \frac{\psi_s x}{l} \right] = \theta_0. \tag{22}$$

We shall presently return to the problem of the determination of the A's, but shall first discuss the forms of the G's and the ψ_s's.

FORMS OF THE G's FOLLOWING FROM EQS. 3 AND 3.1, RESPECTIVELY

We can proceed either by use of (21) or, more simply, by use of (12). Equation 3 requires that

$$Ms\dot{a}_0 = ka_1$$

or, using (5)

$$\dot{a}_0 = \lambda_s l a_1.$$

Using (20)

$$-\sum_0^\infty A_s \lambda_s e^{-\lambda_s t} = -\lambda_s l \sum_0^\infty \frac{A_s R}{l} \frac{\lambda_s}{\lambda_s} G_s e^{-\lambda_s t}.$$

Since this relation must hold for all values of t, it must hold independently for each s. Hence

$$G_s = \frac{1}{R}. \tag{23}$$

Thus the boundary condition embodied in (3) leads to a value of G_s which is independent of s.

Considering now the requirement embodied in (3.1), we have, using (12)

$$Ms\dot{a}_0 - \frac{Msk\dot{a}_1}{h} - ka_1 = 0.$$

Using (5)

$$\dot{a}_0 - \frac{k\dot{a}_1}{h} - \lambda_s l a_1 = 0. \tag{24}$$

Using (20), and observing that (24) must hold for all values of s, individually, we have

$$-\lambda_s = \frac{k}{h} \frac{R\lambda_s^2}{l\lambda_s} G_s - \lambda_s \frac{lR\lambda_s}{l\lambda_s} G_s,$$

$$\left(R + \frac{kR\lambda_s}{hl\lambda_s} \right) G_s = 1 \tag{25}$$

$$G_s = \frac{1}{R(1 + k\lambda_s/hl\lambda_s)},$$

Observing from (18) that $\psi_s = (R\lambda_s/\lambda_s)^{1/2}$, Eq. 25 becomes

$$G_s = \frac{1}{R(1 + k\psi_s^2/Rhl)} \tag{26}$$

so that now G_s does depend upon s.

THE NATURE OF THE ψ_s's AND THE λ_s's

The ψ_s's are determined by (19) which yields, through (23), for the condition of Eq. 3,

$$\psi_s \tan \psi_s = R \tag{27}$$

and through (26) for the condition of Eq. 3.1

$$\psi_s \tan \psi_s = R\left(1 + \frac{k\psi_s^2}{Rhl}\right). \tag{28}$$

The λ_s's are determined immediately in terms of the ψ_s's from (18).

Considering both (27) and (28), it is obvious that the smallest ψ, that is, ψ_0, which we shall call β_0, lies in the first quadrant, and successive β's in the 3rd, 5th, 7th, etc., quadrants. The ψ_s's will therefore be of the form

$$\psi_0 = \beta_0; \quad \psi_1 = \pi + \beta_1; \quad \psi_2 = 2\pi + \beta_2; \quad \psi_s = s\pi + \beta_s$$

with all the β's less than $\pi/2$. Thus (28) assumes the form

$$(s\pi + \beta_s) \tan \beta_s = R\left[1 + \frac{k(s\pi + \beta_s)^2}{Rhl}\right]. \tag{29}$$

Case Corresponding to Equation of Condition (3)—Absence of Thermal Resistance at the Junction of S *and* B

For this case h is infinite, and (29) assumes the form

$$(s\pi + \beta_s) \tan \beta_s = R \tag{30}$$

so that the β_s diminish with increasing s and, in fact, for large s, assume the form

$$\beta_s = \tan \beta_s = \frac{R}{s\pi}.$$

Thus, for large values of s, ψ_s assumes the form

$$\psi_s = s\pi(1 + R/s^2\pi^2) \tag{31}$$

or, to high approximation

$$\psi_s = s\pi \tag{32}$$

for large s.

From (18) and (31)

$$R\lambda_s/\lambda_a = s^2\pi^2(1 + R/s^2\pi^2)^2.$$

Thus, for large s,

$$\lambda_s = \frac{s^2\pi^2}{R}\lambda_a. \tag{33}$$

Even for R as large as 0.5, and $s = 1$, $\lambda_s > 20\lambda_a$. The result is that in our expression (21) for θ, only the term involving the smallest λ_s, which we shall call λ_0, is significant in the sense that if, for example, $e^{-\lambda_0 t}$ is 0.5, $e^{-\lambda_1 t}$ will be less than $(0.5)^{20}$, and so on. Thus, unless the A_s's for large s grow to enormous sizes, which they do not, but, in fact, decrease with increasing s, we need concern ourselves only with the λ_0, except very near to $t = 0$.

It is now of interest to ascertain how near λ_0 is to λ_a. The value of β_0 (corresponding to ψ_0) is given by

$$\beta_0 \tan \beta_0 = R. \tag{34}$$

For the case where R is small, it is necessary to recognize, using (18) that

$$\beta_0(= \psi_0) = (R\lambda_0/\lambda_a)^{1/2} \tag{35}$$

where λ_0 here refers to the smallest of the λ's. Thus

$$\left(\frac{R\lambda_0}{\lambda_a}\right)^{1/2} \tan\left(\frac{R\lambda_0}{\lambda_a}\right)^{1/2} = R.$$

For very small values of R, we can thus write

$$\left(\frac{R\lambda_0}{\lambda_a}\right)^{1/2}\left(\frac{R\lambda_0}{\lambda_a}\right)^{1/2} = R,$$

yielding $\lambda_0/\lambda_a = 1$ for the limiting case where $R = 0$, as might be expected. For the case where R is small but not zero, we may expand the tangent in (34) and write

$$R = \beta_0\left(\beta_0 + \frac{\beta_0^3}{3} + \frac{2\beta_0^5}{15}\right) = \beta_0^2\left(1 + \frac{\beta_0^2}{3} + \frac{2\beta_0^4}{15}\right). \tag{36}$$

The first approximation gives $\beta_0^2 = R$. Substituting this value for β_0^2 in the second parenthesis of (36)

$$\beta_0^2\left(1 + \frac{R}{3} + \frac{R^2}{45}\right) = R.$$

Hence, using (35)

$$\left(\frac{R\lambda_0}{\lambda_a}\right) = \beta_0^2 = \frac{R}{1 + R/3 + R^2/45}$$

$$\frac{\lambda_0}{\lambda_a} = \frac{1}{1 + R/3 + R^2/45}. \tag{37}$$

If $R \sim 0.2$, λ_0 will be less than λ_a by about 7 per cent. The approximate thermal conductivity is determined by writing

$$\lambda_0 = k/Msl,$$

where λ_0 is the experimentally determined value. The more correct expression to yield k in terms of the experimentally measured λ_0 is

$$\lambda_0 = \frac{\lambda_a}{1 + R/3 + R^2/45} = \frac{k}{Msl(1 + R/3 + R^2/45)}. \tag{38}$$

Case Corresponding to Equation of Condition (3.1)—Thermal Resistance at the Junction of S *and* B *Recognized*

We must now use Eq. 28 which leads to (29). It no longer results that the β's necessarily decrease with increase of s. In fact, if h is very small, so that the thermal resistance at the junction is large, (29) assumes the form

$$\tan \beta_s = \frac{k}{hl}(s\pi + \beta_s)$$

so that, with s infinite, $\tan \beta_s$ is infinite and $\beta_s = \pi/2$, which is the largest value of β_s attainable, since it must be in the first quadrant by definition.

If, however, we consider the case where the magnitude of h is such as to permit β_0 to be significantly less than unity, so that we may use the first two terms of the expansion of $\tan \beta_0$, we shall have, from (28),

$$\beta_0\left(\beta_0 + \frac{\beta_0^3}{3}\right) = \beta_0^2\left(1 + \frac{\beta_0^2}{3}\right) = R + \frac{k}{hl}\beta_0^2.$$

Our first approximation to β_0^2 is given by

$$\beta_0^2 = \frac{R}{1 - k/hl}.$$

Substituting this for β_0^2 in the second parenthesis, and neglecting terms

involving products of R and k/hl, we readily find

$$\beta_0^2 = \frac{R}{\left(1 - \frac{k}{hl} + \frac{R}{3}\right)}$$

so that, from (35)

$$\frac{\lambda_0}{\lambda_\infty} = \frac{1}{\left(1 - \frac{k}{hl} + \frac{R}{3}\right)}.$$

We see, as might be expected, that the effect of the thermal resistance at the junction is to diminish the difference between λ and λ_0.

Again, if λ is the experimentally measured value, the expansion for determining the conductivity k should be

$$\frac{k}{Msl} = \left(1 - \frac{k}{hl} + \frac{R}{3}\right)\lambda_0.$$

In principle it is possible to determine h by performing two experiments with different values of l. In this way, two values of λ_0 are experimentally measured, and provide us with two equations from which k and h may be obtained.

While expansion of $\tan \beta_0$ has been invoked in the foregoing discussions of both cases, that where thermal resistance is ignored and that where it is included, the use of such expansions is unnecessary. We can always handle the equations concerned graphically. We have used the expansions mainly to provide ease of description.

THE DETERMINATION OF THE A_n's IN EQ. 21

While the complete solution for θ as a function of x and t is not necessary for the practical purpose of determining the relation which governs the measurement of k, it is of general interest to complete the solution expressed in (21) by outlining the processes which may be used for the determination of the A_n's. The fundamental equation for this purpose is (22).

Of course, we can always adopt the procedure of starting by limiting the infinite sum on the left hand side of (22) to a finite sum of N terms, writing down N equations to provide a fit of the series to N points from $x = 0$ to $x = l$. We then have N linear equations to determine the corresponding N coefficients A_n, the results, of course, being only approximations. By repeating the process with a larger number of terms, we can trace the progress of the approximation to any degree we desire; but this process is laborious, and unimaginative for the purpose of bringing out the controlling influences of the various magnitudes involved. However, other methods of determining the A_n's present peculiar difficulties in the present instance.

If the expression on the left hand side of (22) were in the form of a Fourier series, with independent constants for the sine and cosine terms, the constants to be determined so as to produce the right hand side, all would be straightforward and simple as regards the determination of the constants. However, in the first place the constants associated with the sine and cosine terms are bound by a common A_s. Moreover, on account of the rather complicated nature of the β's occurring in the ψ_s's, the coefficients of the A_s's are not even "normal functions." For this reason, our determination of the A_s's must follow a line which is dictated by these circumstances.

Recalling our discussion on the form of the G_s's, we infer that, for any boundary condition at $x = 0$, there is an infinite set of λ's, denoted by the λ_s's, and an infinite set of G_s's related to the λ_s's by the conditions which, in line with (19), are

$$\psi_s \tan \psi_s = 1/G_s \tag{39}$$

where, in line with (18)

$$\psi_s = (R\lambda_s/\lambda_a)^{1/2}. \tag{40}$$

With this understanding, we can write (22) as

$$\sum_0^\infty A_s \left[\cos \psi_s \frac{x}{l} - \cot \psi_s \sin \psi_s \frac{x}{l}\right] = \theta_0$$

that is,

$$\sum_0^\infty \frac{A_s}{\sin \psi_s}\left[\cos \psi_s \frac{x}{l} \sin \psi_s - \sin \psi_s \frac{x}{l} \cos \psi_s\right] = \theta_0$$

$$-\sum_0^\infty \frac{A_s}{\sin \psi_s}\left[\sin \psi_s \left(\frac{x}{l} - 1\right)\right] = \theta_0. \tag{41}$$

Let us multiply by $\sin \psi_m \left(\frac{x}{l} - 1\right)$ and integrate from $x = 0$ to $x = l$

$$\sum_0^\infty \frac{A_s}{\sin \psi_s} \int_0^l \sin \psi_s \left(\frac{x}{l} - 1\right) \sin \psi_m \left(\frac{x}{l} - 1\right) dx = -\theta_0 \int_0^l \sin \psi_m \left(\frac{x}{l} - 1\right) dx$$

$$-\sum_0^\infty \frac{Al}{2 \sin \psi_s} \int_0^l \cos(\psi_m + \psi_s)\left(\frac{x}{l} - 1\right) - \cos(\psi_m - \psi_s)\left(\frac{x}{l} - 1\right) = \frac{l\theta_0}{\psi_m}\left[\cos \psi_m \left(\frac{x}{l} - 1\right)\right]_0^l$$

$$-\sum_0^\infty \frac{Al}{2 \sin \psi_s}\left[\frac{\sin(\psi_m + \psi_s)}{\psi_m + \psi_s} - \frac{\sin(\psi_m - \psi_s)}{\psi_m - \psi_s}\right] = \frac{l\theta_0}{\psi_m}[1 - \cos \psi_m]$$

$$-\sum_0^\infty \frac{Al}{2 \sin \psi_s}\left[\frac{\sin \psi_m \cos \psi_s + \cos \psi_m \sin \psi_s}{\psi_m + \psi_s} - \frac{\sin \psi_m \cos \psi_s - \cos \psi_m \sin \psi_s}{\psi_m - \psi_s}\right] = \frac{l\theta_0}{\psi_m}(1 - \cos \psi_m)$$

$$-\sum_0^\infty \frac{lA_s \sin \psi_m}{2 \sin \psi_s} \sin \psi_s \left[\frac{\cot \psi_m + \cot \psi_s}{\psi_m + \psi_s} + \frac{\cot \psi_m - \cot \psi_s}{\psi_m - \psi_s}\right] = \frac{\theta_0 l}{\psi_m}(1 - \cos \psi_m).$$

Using (39)

$$-\sum_0^\infty \frac{lA_s}{2} \sin \psi_m \left[\frac{G_m\psi_m + G_s\psi_s}{\psi_m + \psi_s} + \frac{G_m\psi_m - G_s\psi_s}{\psi_m - \psi_s}\right] = \frac{\theta_0 l}{\psi_m}(1 - \cos \psi_m). \tag{42}$$

For an assigned m, this expression holds with the sum extending to all A_s from $s = 0$ to ∞, except that the contribution from A_m to the left hand side should be specially evaluated in view of a certain indeterminancy in the term involving $\psi_m - \psi_s$ in the denominator. To this end we omit the term involving A_m on the left hand side of (42) and replace it by J_m, where

$$J_m = \frac{A_m}{\sin \psi_m}\left[\int_0^x \sin^2 \psi_m \left(\frac{x}{l} - 1\right)\right] dx$$

which is the appropriate contribution, as may be seen by (41). We thus have

$$J_m = \frac{-A_m}{2 \sin \psi_m} \int_0^l \left[\cos 2\psi_m \left(\frac{x}{l} - 1\right) - 1\right] dx$$

$$= \frac{-A_m l}{2 \sin \psi_m}\left[\frac{\sin 2\psi_m}{2\psi_m}\left(\frac{x}{l} - 1\right) - x\right]_0^l = \frac{-A_m l}{2 \sin \psi_m}\left[\frac{\sin 2\psi_m}{2\psi_m} - 1\right]$$

$$J_m = \frac{-A_m l}{2}\left[\sin \psi_m \frac{\cot \psi_m}{\psi_m} - \frac{1}{\sin \psi_m}\right] = \frac{-A_m l}{2}\left[G_m \sin \psi_m - \frac{1}{\sin \psi_m}\right].$$

Equation 42 thus yields

$$-\sum_0^{m-1} \frac{lA_s}{2} \sin \psi_m \left[\frac{G_m\psi_m + G_s\psi_s}{\psi_m + \psi_s} + \frac{G_m\psi_m - G_s\psi_s}{\psi_m - \psi_s}\right]$$

$$-\sum_{m+1}^{\infty} \frac{lA_s}{2} \sin \psi_m \left[\frac{G_m\psi_m + G_s\psi_s}{\psi_m + \psi_s} + \frac{G_m\psi_m - G_s\psi_s}{\psi_m - \psi_s}\right]$$

$$-\frac{lA_m}{2} G_m \sin \psi_m + \frac{lA_m}{2 \sin \psi_m} = \frac{l\theta_0}{\psi_m}(1 - \cos \psi_m). \qquad (43)$$

Case Where Thermal Resistance at x = 0 *is Neglected*

In this case, we observe that G_s is independent of s and is equal to $1/R$. Thus for this case, (43) becomes

$$-\sum_0^{m-1} \frac{A_s}{R} \sin \psi_m - \sum_{m+1}^{\infty} \frac{A_s}{R} \sin \psi_m + \frac{A_m}{2R} \sin \psi_m + \frac{A_m}{2 \sin \psi_m} = \frac{\theta_0}{\psi_m}(1 - \cos \psi_m)$$

$$-\sum_0^{\infty} \frac{A_s}{R} \sin \psi_m + \frac{A_m}{2R}\left(\sin \psi_m + \frac{R}{\sin \psi_m}\right) = \frac{\theta_0}{\psi_m}(1 - \cos \psi_m) \qquad (44)$$

$$-\sum_0^{\infty} A_s = \frac{R\theta_0}{\psi_m}\left(\frac{1 - \cos \psi_m}{\sin \psi_m}\right) + \frac{A_m}{2}\left(1 + \frac{R}{\sin^2 \psi_m}\right).$$

Again, using (39), this becomes

$$-\sum_0^{\infty} A_s = \frac{\theta_0}{\cos \psi_m} - \theta_0 - \frac{A_m}{2}\left(1 + \frac{R}{\sin^2 \psi_m}\right). \qquad (45)$$

Now referring to (41), let us apply the result at $x = 0$, in which case it yields

$$\sum_0^\infty A_s = \theta_0. \tag{46}$$

Thus, (45) yields

$$A_m = \frac{2\theta_0}{\cos \psi_m (1 + R/\sin^2 \psi_m)}.$$

At this stage we may change the subscript m back to s, and write

$$A_s = \frac{2\theta_0}{\cos \psi_s (1 + R/\sin^2 \psi_s)}.$$

Recalling that $\psi_s = s\pi + \beta_s$, we readily see that $\cos \psi_s = (-1)^s \beta_s$, and $\sin^2 \psi_s = \sin^2 \beta_s$. Thus

$$A_s = \frac{2(-1)^s \theta_0}{\cos \beta_s (1 + R/\sin^2 \beta_s)}. \tag{47}$$

The substitution of (47) in (21), with G_s replaced by $1/R$ for the case under consideration, completes the solution of our problem for the determination of θ as a function of x and t in a form satisfying all the conditions $A - E$. Referring to the development between Eqs. 39 and 41, we see that, for the case under consideration,

$$\sum_0^\infty A_s \left[\cos \psi_s \frac{x}{l} - G_s \psi_s \sin \frac{x}{l}\right] = -\sum_0^\infty \frac{A_s}{\sin \psi_s} \sin \psi_s \left(\frac{x}{l} - 1\right).$$

Thus, since $\sin \psi_s = \sin (s\pi + \beta_s) = (-1)^s \sin \beta_s$

$$\theta = \sum_0^\infty \frac{2\theta_0 e^{-\lambda_s t} \sin \psi_s (1 - x/l)}{\sin \beta_s \cos \beta_s (1 + R/\sin^2 \beta_s)}. \tag{48}$$

Comment Concerning the A_s*'s*

While the derivation of the A_s's as given in (47) seems perfectly logical, the procedure adopted is, as far as the writer is aware, rather unusual, and there would seem to be some mental satisfaction in applying some independent test. Such a test would be to show that, with the A_s's as given by (47), $\sum_0^\infty A_s = \theta_0$ as demanded by (46). The application of this test seems to involve rather cumbersome procedures for the general case. However, it is possible to appy the test without much difficulty to the case where R is small. This we shall do.

Our basic Eq. 27 yields (30), which is

$$(s\pi + \beta_s) \tan \beta_s = R. \tag{49}$$

For $s = 0$, this yields

$$\beta_0 \tan \beta_0 = R.$$

For R small, both β_0 and $\tan \beta_0$ are small, so that we may write

$$\beta_0 \left(\beta_0 + \frac{\beta_0^2}{3}\right) = \beta_0^2 \left(1 + \frac{\beta_0^2}{3}\right) = R.$$

Substituting the first approximation $\beta_0^2 = R$ in the second parenthesis, we have

$$\beta_0^2 = \frac{R}{1 + R/3}$$

$$\sin \beta_0 = \beta_0 - \frac{\beta^3}{3} = \beta_0 \left(1 - \frac{\beta_0^2}{3}\right)$$

$$\sin^2 \beta_0 = \frac{R}{(1 + R/3)} \left(1 - \frac{2R}{3}\right)$$

$$\frac{R}{\sin^2 \beta_0} = 1 + \frac{2R}{3} \qquad \text{to the first power of } R$$

$$\cos \beta_0 = 1 - \frac{\beta_0^2}{2} = 1 - \frac{R}{2} \qquad \text{to the first power of } R.$$

Hence, from (47)

$$A_0 = \frac{2\theta_0}{(1 - R/2)(2 + 2R/3)}.$$

Retaining only first powers of R, we have

$$A_0 = \theta_0 \left(1 + \frac{R}{6}\right). \tag{50}$$

For $s > 0$, Eq. 49 yields as a first approximation

$$\tan \beta_s = \beta_s = \frac{R}{s\pi}.$$

Writing (49) as

$$(s\pi + \beta_s)\beta_s \left(1 + \frac{\beta_s^2}{3}\right) = R$$

and substituting the first approximation for β_s in the two parentheses,

we have

$$s\pi\left(1+\frac{R}{s^2\pi^2}\right)\left(1+\frac{R^2}{3s^2\pi^2}\right)\beta_s = R,$$

so that, retaining only the first power of R, we have

$$\beta_s = R/s\pi$$

and to the same order $\sin^2\beta_s = \beta_s^2 = R^2/s^2\pi^2$, and

$$\frac{R}{\sin^2\beta_s} = \frac{s^2\pi^2}{R}.$$

Again,

$$\cos\beta_s = 1 - \frac{\beta_s^2}{2} = 1 - \frac{R^2}{2s^2\pi^2}.$$

Thus, from (47) we have, for $s \geqslant 1$,

$$A_s = \frac{2(-1)^s\theta_0}{(1-R^2/2s^2\pi^2)\left(1+\frac{s^2\pi^2}{R}\right)}$$

$$= \frac{2(-1)^sR\theta_0}{s^2\pi^2\left(1-\frac{R^2}{2s^2\pi^2}\right)\left(1+\frac{R}{s^2\pi^2}\right)}.$$

Retaining only the first power of R

$$A_s = \frac{2(-1)^sR}{s^2\pi^2}\theta_0.$$

Thus

$$\sum_1^\infty A_s = R\sum_1^\infty\frac{2(-1)^s}{s^2\pi^2}\theta_0 = -2R\left(\frac{1}{12}\right)\theta_0 = -\frac{R}{6}\theta_0. \qquad (51)$$

Now

$$\sum_0^\infty A_s = A_0 + \sum_1^\infty A_s.$$

So that, using (50) and (51)

$$\sum_0^\infty A_s = \left(1+\frac{R}{6}\right)\theta_0 - \frac{R}{6}\theta_0 = \theta_0.$$

It is not intended to imply that (46) holds for *only* small values of R. Our development requires, as a logical necessity, that it holds for all values of R. However, since it is not easy to verify this fact directly and independently, starting with (47), we have carried out this test for the algebraically simple cases where R is small.

NOTE CONCERNING THE EFFECT OF THERMAL JUNCTION RESISTANCE

The conductivity is calculated from λ_0 which, to the approximation involved in retaining only the λ_0 term in the solution for θ_b at the junction, gives $\theta_b = \theta_0 e^{-\lambda_0 t}$. The measured boundary temperature θ_B differs from θ_b by $\Delta\theta$ cited in footnote (2). However, the relations in footnote (2) readily yield

$$\theta_B = \theta_0 e^{-\lambda_0 t} \Big/ \left(1 - \frac{\lambda_0 Ms}{h}\right) = \theta_b/(1 - \lambda_0 Ms/h).$$

Thus the slopes of $\log \theta_B$ and $\log \theta_b$ as functions of t are the same, so that use of θ_B instead of θ_b does not affect the determination of λ_0.

THE HUMAN OPERATOR AS A SERVO SYSTEM ELEMENT*

BY

DUANE T. McRUER[1] AND EZRA S. KRENDEL[2]

PART I**

SUMMARY

This paper considers the role of human elements in certain closed loop control systems. A quantitative description of human dynamics useful to control system designers is essential to understanding and analysis of such systems. Accordingly, the human behavior description must be expressed in terms which are compatible with conventional descriptions of control system components. This compatibility is achieved by the use of a quasi-linear mathematical model for the human operator. The model is composed of two components—a describing function and remnant. The describing function, which for a linear system is identical with the conventional transfer function, is established to characterize that portion of the operator's output which is linearly correlated with his input. The input, upon which the describing function is based, is selected on the basis of *a priori* estimates of the nature of certain human nonlinear behavior. Human output power which cannot be characterized by the operation of the describing function on the input is designated as the remnant.

After presenting the analytic basis for measurements of human dynamics, steady-state describing functions measured by the various experimenters in the field are discussed and the adaptive, optimalizing behavior of the human operator is demonstrated. The remnants are also discussed and plausible sources for their origin are postulated.

Knowledge of the range of parameter adjustment of which the human operator is capable in his adaptation as well as knowledge of his criteria for adjusting these parameters enables the designer to specify input functions and operator controlled dynamics compatible with both desirable human operator behavior and good system performance. By judiciously trading off system complexity against operator preferences, while still making proper engineering use of the human operator's adaptability, a control system may be optimized for both performance and reliability.

INTRODUCTION

The control engineer achieves a logical connection between a system and its components by expressing component characteristics in terms of abstract models, such as transfer functions or computer setups, and then combining these models to obtain a description of system behavior. The use of such techniques in the design of automatic control systems has met with outstanding success. In this process, designers have tended to overlook one of the most versatile and, in many respects, most

* This research was supported in whole or in part by the United States Air Force under Contracts AF 33(616)–3080, AF 33(038)–10420, AF 33(616)–2804, AF 33(616)–3610, and AF 33(616)–5822, monitored jointly by the Flight Control Laboratory and the Aero Medical Laboratory, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

[1] Systems Technology, Inc., Inglewood, Calif.

[2] The Franklin Institute Laboratories for Research and Development, Philadelphia, Pa.

** Part II will appear in this JOURNAL for June, 1959.

readily available and reliable components of all—the human controller. There are fairly broad areas where a suitable manual control system represents either the optimum solution or merits detailed consideration as a good possible choice. Before control engineering concepts can be applied to such cases, however, one must have a catalog of human static and dynamic characteristics, preferably expressed in terms readily understood by the control engineer. One of the principal purposes of this paper is to provide the reader with a synthesis, in control engineering terminology, of the available data on the dynamic characteristics of human operators in certain continuous control tasks. This summary is to a large extent based upon McRuer and Krendel (1).[3] Much of the available static characteristic data can be found in (2 and 3).

Many agencies and individuals, both in the U.S.A. and abroad, have made substantial contributions to the state of knowledge of human response characteristics. In the past these data have been compared rarely if ever, since they were measured in different ways, different tasks were given the operator, diverse system inputs were utilized, and the data were presented in different forms. A review of the situation in the light of servo analysis indicates that consistencies exist for these data, from which useful generalizations can be hypothesized about the human operator dynamics.

V.2
195

The human response measurements of concern here are those involved in closed loop control tasks which yield operator transfer characteristics for routine, essentially continuous functions such as stabilization and tracking. Even for these simple tasks we cannot expect to obtain a unique transfer characteristic since the operator's responses will depend upon at least the following factors:

1. The dynamic characteristics of the controlled elements, that is, the dynamics between the manipulated variable and the display. (The operator, to be successful, must adapt his transfer characteristics to a form required for adequate stability and performance of the entire closed loop system.)
2. The type of input or forcing function driving the system.
3. The individual reaction delays, thresholds, etc., of the human during the particular operation. (There are variations between individuals and within individuals which are functions of both physiological differences and task experience. Any set of transfer characteristics ultimately proposed would have to allow for a variation in parameters of this type.)
4. The motivation, attention, previous training, and general psychological and physiological condition of the human at the time of the operation.

These considerations specify a human mechanism which adapts itself

[3] The boldface numbers in parentheses refer to the references appended to Part I of this paper.

in some way as a function of the system inputs and controlled elements, and which may have wide individual variations in the values of the parameters adopted. Individual parameter variations cannot be specified at present since there are too few data dealing with adequately large populations. There do exist adequate data to demonstrate the adaptability of human transfer characteristics to both forcing functions and controlled elements and such data provide a rationale for classifying past experiments.

The selection of the form of the mathematical model to describe the operator's transfer characteristic is based upon several practical considerations: the ease with which it lends itself to conventional servo analysis methods; the convenience with which experimental data may be obtained to define such a model; and the extent to which the measuring experiments are capable of generalization. These factors limit human operator models to the following:

1. A linearized model consisting of an equivalent quasi-linear transfer function (which is correlated linearly with the system input or forcing function) and a remnant or residual output (which represents all output content which cannot be ascribed to a linear operation on the input).
2. Simulation, using analog computer elements.
3. A series of nonlinear or piece-wise linear equations which can be solved by either analog or digital means.

V.267
1959

To date only the first two have been used successfully as the basis of experimental techniques. The observation that human operator control responses are discontinuous has led to the conjecture that a nonlinear relay servo model would provide a more natural characterization for human control behavior. Although certain preliminary experimental results indicated the existence of possible dual or multi-mode relay type servo behavior on the part of the human, other data shed doubt on this behavior (4, 5). The literature demonstrates that the model lacks convenient applicability to the analysis of continuous closed loop systems with continuous inputs (6, 7).

To provide an adequate representation of the operator in a closed loop controller application, the operator mathematical models should provide answers to two questions:

1. How does (will) the operator perform in a given situation, that is, what specific quasi-linear transfer function and remnant, or analog computer set-up, are required to describe his conduct?
2. How well can the operator perform and how do we make him behave that way, that is, what is the "desirable" operator characteristic and under what circumstances is it achieved?

The answer to the first of these questions is essential for manual control system analysis, while that of the second provides the starting point for

synthesis. The first question is answered by the formulation of an adaptive-optimalizing model developed to be consistent with all of the available data. This model can be used, with caution, for stability and limited performance predictions of manual control systems. The second question is answered to the extent of defining a desirable system form and the human operator's response when operating with this form. This "desirable" operator characteristic refers to a "best" characteristic within the adaptive properties of a so-called "average" human operator, and *not* to an ideal super-human-machine combination. What is desirable from the system designer's lights may not be preferred from the controller's viewpoint.

In developing the answers to the above questions this paper is organized in four main parts. The first is primarily background information on the types of manual control systems of interest to this paper, and the general methods for the description of quasi-linear operation. The second presents a summary of presently available human dynamic response data, gathered from experiments directly involving the measurement of human describing functions from tracking systems which display the error alone. Past experimental efforts are reviewed, and the approximate mathematical models evolved to "fit" the data are presented. The third and fourth parts generalize from the experimental data.

The hypothetical models discussed are in two categories: a general hypothetical model attempting a mathematical description of the human operator for various system considerations (that is, various controlled element configurations and types of system inputs); and a "desirable" human operator form determined by a minimum rms error criterion of system performance. Since the form of the highly adaptive human operator is determined by the control problem, the controlled elements of the system can be subjected to design change to elicit desirable human operator characteristics. By having both a general hypothetical operator model and a preferred operator model, the control system analyst can evaluate the performances of particular system configurations with that of a more or less theoretical optimum, and can gain an idea of the type of modification required to the rest of the system if system error performance alone were the sole design objective. Since other factors such as reliability, maintainability and cost are often as important as dynamic performance, the designer must effect a compromise in order to achieve an over-all optimum.

TYPES OF MANUAL CONTROL SYSTEMS AND METHODS OF OPERATOR DESCRIPTION

Operator Environmental Factors Due to the Control System

The fact of human adaptation makes description of the operator enormously complex when viewed in the large, and necessitates the

setup of simplified, constant situations within which one may have some hope of obtaining a simple behavioral model of engineering value.

For the control situation to be analyzed with confidence the assigned task or goals of the human controller, and the constraints or rules governing his interactions with the system, must remain reasonably constant over the duration of interest. In the present instance we are concerned with manual processes for minimizing visually perceived errors by exercising essentially continuous control so as to match visually presented input and output signals. This type of process is often called tracking.

One can classify tracking problems for conceptual convenience. Our criteria for classification will be based on the type of information which the stimulus to be tracked and the controller's response present to the controller as a basis for future tracking decisions. In application, these criteria result in three limiting types of classifications: precognitive, pursuit and compensatory. These classifications arise as interactions among training, the geometry of the display, and the nature of the input. As such, they can readily exist in mixed forms.

Precognitive. This condition exists when the operator has complete information about the input's future and a stimulus can trigger off a repertory of practiced, properly sequenced responses. In a sense, the situation itself may be the stimulus. Thus throwing a baseball at a target is precognitive behavior, as is steering a car out of a skid, or navigational flying under VFR conditions. The operator doesn't need to maintain a frequent check on the individual responses in a sequence. Instead, end product responses are monitored; such as, "Did I hit the target?", "Am I still skidding?" or "Have I passed over a given fix?" In that continuous close control *is not* maintained on the perceived error, precognitive behavior is not tracking. However, we have included this classification because occasionally tracking approaches these conditions. One might characterize precognitive behavior as discrete control with long sampling intervals such that open loop control is in effect during these intervals.

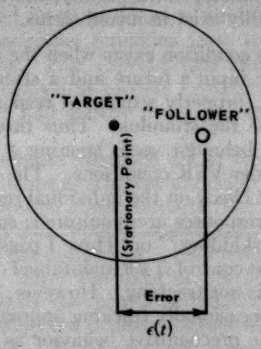
Pursuit. In pursuit behavior past experience provides the tracker with information with which to predict the future input, but he must operate in a closed loop fashion with visual feedback about his responses. In this form of tracking the display effects of the operator's corrective responses can be distinguished from his input. These responses are displayed *after* they have been modified by the system dynamics. A gunner aiming for a moving target using an open sight is an example of pursuit tracking in which the system dynamics are those of the mechanical system composed of gun, arm, and body.

Compensatory. Compensatory tracking is the same as pursuit except that the visually displayed effects of the controller's responses are not distinguishable from the system's input. Were our gunner aiming for

the target using a telescopic sight with a small field of view, we would have a compensatory situation. Since in compensatory tracking the visual display is the system forcing function minus the modified control response, the operator can determine the effects of control motion alone only under zero input conditions.

The precognitive classification, representing a skilled operator response or calibrated, open loop system, is not readily amenable to servo analysis techniques. The pursuit system is a particular type of open cycle, closed cycle servo system, and requires a more complicated servo analysis (8). The compensatory classification, representing continuous close-control tasks, can be attacked by simple servo methods. The particular type of manual control system to be considered here falls into the compensatory category.

The general nature of a compensatory display is represented in Fig. 1. The operator is presented with an input consisting only of an



$$\epsilon(t) = i(t) - r(t)$$

FIG. 1. Pure compensatory display.

indicator showing the difference, or error, $\epsilon(t)$, between a system forcing function, or command signal, $i(t)$, and the system output $r(t)$. The operator's task is to minimize the error signal presented by trying to keep the circle superimposed on the stationary dot.

Besides the display, other fundamental factors in the control situation are the type of element being controlled by the operator (such as an aircraft, automobile, etc.), and the actual means of exerting control (such as a control stick or wheel with their associated restraints, for example, springs, dampers, etc.). All of these characteristics will be lumped into the general classification of the "controlled element." In many cases the controlled element can be described suitably by transfer

functions, though in others nonlinearities may be present and require description.

The functional block diagrams of the simple control situation considered, with the human operator as an element of a closed loop system, is shown in Fig. 2.

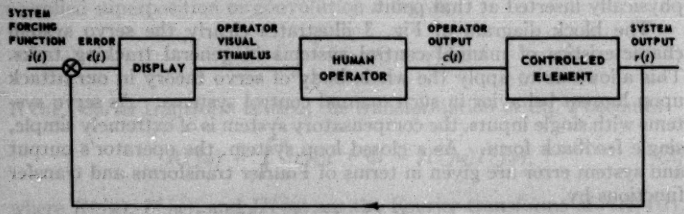


FIG. 2. Functional block diagram of simple compensatory manual control system.

Quasi-Linear Operator Description

If the characteristics of the human operator for a given over-all task are assumed to be capable of quasi-linear description, the operator mathematical model will consist of a describing function plus an additional quantity inserted as an input into the system by the operator. Then the functional block diagram of Fig. 2 can be made into an equivalent block diagram showing more detailed operations of the system. This is done in Fig. 3, where the symbols describe the various quantities

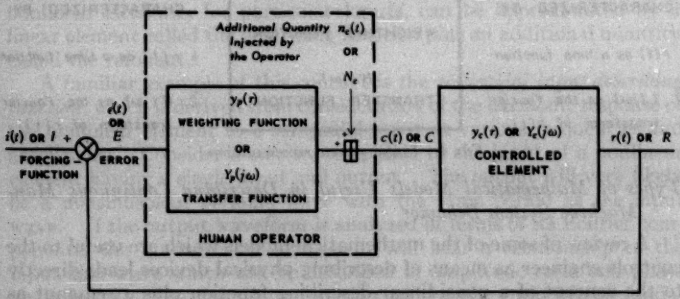


FIG. 3. Block diagram showing human operator in continuous control task.

as time functions and as Fourier transforms. To simplify the structure, the dynamic characteristics of the display are lumped with the controlled element, and the actual forcing function is modified, if necessary, to an equivalent one. The linear transfer characteristics of the human operating on his presented input are described by the weighting function, $y_p(\tau)$, or the Fourier transform, $Y_p(j\omega)$. Since all of the operator's

output is not described by the action of a linear transfer characteristic, an additional term, $n_e(t)$, is added at his linear output to complete the description of the total operator output $c(t)$. This additional term is called the remnant. The location of $n_e(t)$ at the operator's output is arbitrary, and does not necessarily imply that such a quantity is physically inserted at that point.

The block diagram of Fig. 3 illustrates clearly the servo system characteristics of manual control systems in general tracking tasks. This allows us to apply the whole body of servo theory in our attack upon human behavior in such manual control systems. As servo systems with single inputs, the compensatory system is of extremely simple, single feedback form. As a closed loop system, the operator's output and system error are given in terms of Fourier transforms and transfer functions by,

$$C(j\omega) = \frac{Y_p(j\omega)I(j\omega) + N_e(j\omega)}{1 + Y_p(j\omega)Y_e(j\omega)}$$

$$E(j\omega) = \frac{I(j\omega) - Y_e(j\omega)N_e(j\omega)}{1 + Y_p(j\omega)Y_e(j\omega)} \qquad (1)$$

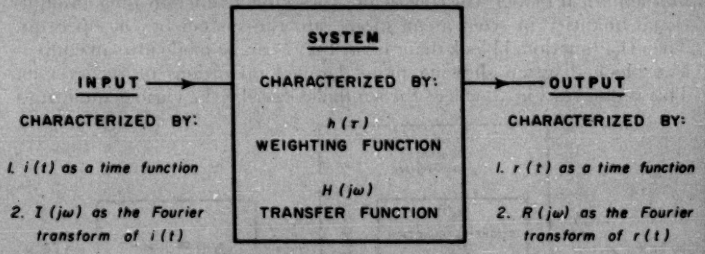


Fig. 4. Linear system representation.

Types of Mathematical Models Useful in Describing Continuous Man-Machine System Behavior

A review of some of the mathematical models which are useful to the controls engineer as means of describing physical devices leads directly to the concept of a quasi-linear describing function plus a remnant as basic tools for describing human dynamic characteristics. This concept allows the retention of the linear methods so successfully employed in servo analytical techniques. The remnant so developed is seen from Eq. 1 to be an element in the system error and output equations. In many manual control systems, it can be the dominant operator output quantity in the generation of tracking errors.

With physical systems containing only elements which behave in a way describable by linear constant-coefficient differential equations, the

appropriate mathematical model is the element's weighting or transfer function. If the system were represented by the block diagram of Fig. 4, where the weighting function, $h(\tau)$, is the time response of the system when an impulse function is applied at zero time, then the relationship between the response and the input would be given by the so-called superposition or convolution integral,

$$r(t) = \int_{-\infty}^{\infty} h(\tau)i(t - \tau)d\tau. \tag{2}$$

If the Fourier transform is used, the transformation of Eq. 2 becomes

$$R(j\omega) = \int_{-\infty}^{\infty} r(t)e^{-j\omega t}dt = H(j\omega)I(j\omega), \tag{3}$$

where $R(j\omega)$, $I(j\omega)$, and $H(j\omega)$ are the Fourier transforms of $r(t)$, $i(t)$, and $h(\tau)$, respectively. Laplace transforms can, of course, be used in a similar way.

The introduction of the weighting function, or its transform, the transfer function, allows us to describe the performance of a *linear* system element completely. If the convolution integral is not valid, and the behavior of the system is a function of the *particular* inputs and initial conditions the analyst must define system operation by input-response pairs. Fortunately many nonlinear systems of interest have specific input-specific response pairs which appear to be very similar to input-response pairs for linear systems. Thus the performance of some nonlinear elements, for particular inputs, can be approximated by a linear element called the describing function plus an additional quantity called the remnant.[4]

A familiar example of this concept is the *sinusoidal input describing function*, which is derived from consideration of the harmonic response of the nonlinear element to a sinusoidal input at various frequencies and amplitudes. Consider a sine wave applied to the input of a nonlinear element having a single input and output. The output will very likely be a nonsinusoidal periodic wave with the same period as the input wave. If the output waveform is analyzed in terms of its Fourier components, the fundamental component will bear a relationship to the input sine wave which can be described in amplitude ratio and phase angle terms. The describing function will be the ratio of the fundamental to the input in the same way as used in a linear system. The remnant will be composed of all the higher harmonics.

Since the describing function concept is based upon a particular input type, describing functions can be defined for a particular nonlinear element simply by considering different types of inputs. The

[4] While the notion of molding a nonlinearity into an equivalent linear element is quite old and has been used by many writers, the first instance of major exploitation of the technique was probably by Kryloff and Bogoliuboff (9).

sinusoidal input describing function already mentioned is the most widely known example. Where the type of transient input is known and simply described in analytic form, a transient describing function may be used; for example, the step input describing function. In cases where system input forms may not be simple steps or sine waves, a more useful describing function may be one based upon statistical inputs (10, 11). Such a describing function, particularly for inputs characterized by Gaussian amplitude distributions, is useful in nonlinear control problems where the inputs expected in the application of the control system are statistical, and can be considered as a basis for a fairly general type of equivalent linearization. The Gaussian input describing function will be shown here to be of particular value in describing human response.

As a general comment on describing functions, it should be noted that, since the inputs are different, the only thing the various describing functions have in common is the nonlinear element they intend to approximate. As the elements become less nonlinear, the various describing functions for a given element tend to approach one another. When the nonlinearities are entirely removed, all of the describing functions become identical to the element's transfer function.

Alternate approaches to the nonlinear problem exist. The most practical alternate for human response measurements is time domain analysis by the use of an operator analog comprised of analog computer elements. Using the very wide range of nonlinear and linear analog elements available, a computer setup can be made by cut-and-try procedures and adjusted until the "analog operator" responses to particular inputs are similar to those of the actual operator. If an all-comprehending analog model were ever achieved, the input type would not be important. To achieve a practical and reasonably simple computer setup in the case of the human operator, however, a separate analog is required for different inputs. The analog model technique is of great value in instances where a point by point prediction of operator response is desired, for studies including nonlinear control effects, and as a means of providing insight into some of the types of nonlinear behavior which might occur in the operator. When nonlinear or time-varying transfer characteristics become exceedingly important it is probably the only practical approach. The analog technique is fairly straightforward, so no further discussion of the general method need be included here.

Quantities Involved in Human Response Measurements

Human operator characteristics are strong functions of the type of system input, so the describing function concept is necessary if we wish to retain and apply linear methods of analysis to manual control systems. Almost all experimental research to specify operator dynamics in manual control tasks has been based on this approach. A key factor in the

determination of the operator's response characteristics is the general predictability of the system input. Specifically, the operator's mode of response with either a unity controlled element or a thoroughly learned control element is different for single sine waves (which are essentially completely predictable once recognized), step functions (which have known final values immediately after initiation of the stimulus), and "random appearing functions." The random appearing inputs are not restricted to functions which are known only by their statistics, but include those which can be made up by as few as three nonharmonically related sine waves. In most of the experimental work, the random appearing inputs have been either such a sum of sine waves, or time functions which had essentially Gaussian amplitude distributions. Therefore the types of describing functions of general interest in human response measurement are those for step, sinusoidal, and Gaussian inputs. All of these describing functions fall into a category which can be called quasi-linear because they tend to be linear under a fixed set of conditions, such as given inputs, yet nonlinear when changes in these conditions are considered. Since the statistical forms are usually the closest to reality in practical control systems, we shall confine our attention to transfer characteristics derived for random appearing inputs. When random inputs are used, the most common techniques of describing the characteristics of the time stationary inputs and outputs are by means of correlation functions or power spectral densities. The describing functions are obtained from cross-correlation or cross-spectral density functions, and the remnant and system forcing functions are normally expressed as spectral densities. While most of the underlying theory on the measurement and interpretation of these quantities is beyond our present scope (see, for example, 12 for details), it is necessary to develop some relationship among the fundamental quantities which are used to measure and define the operator's response characteristics.

The various quantities in the compensatory system of Fig. 3 which are important in the determination of operator characteristics are given as a nomenclature summary, in the accompanying tabulation.

The operator's output and the system error for the compensatory system were previously given in Eq. 1. The cross-spectral density functions Φ_{ic} and Φ_{ii} are defined as follows (where the asterisk indicates a conjugate),

$$\Phi_{ic} = \lim_{T\to\infty} \frac{1}{T}[I^*C] = \lim_{T\to\infty} \frac{1}{T}[I(-j\omega)C(j\omega)] \tag{4}$$

$$\Phi_{ii} = \lim_{T\to\infty} \frac{1}{T}[I^*E] = \lim_{T\to\infty} \frac{1}{T}[I(-j\omega)E(j\omega)]. \tag{5}$$

Quantity	Expressed as a Time Function	Expressed as a Fourier Transform*	Expressed in Spectral Form**
System forcing function	$i(t)$	$I(j\omega)$ or I	Φ_{ii}
Error signal of the system; input to the operator	$\epsilon(t)$	$E(j\omega)$ or E	$\Phi_{\epsilon\epsilon}$
Operator output	$c(t)$	$C(j\omega)$ or C	Φ_{cc}
System output response	$r(t)$	$R(j\omega)$ or R	Φ_{rr}
Remnant at the operator's output	$n_e(t)$	$N_e(j\omega)$ or N_e	Φ_{nn_e}
Operator's weighting and describing functions	$y_p(\tau)$	$Y_p(j\omega)$ or Y_p	
Controlled element weighting and transfer functions	$y_c(\tau)$	$Y_c(j\omega)$ or Y_c	
Closed loop weighting and describing function between forcing function and operator's output	$h(\tau)$	$H(j\omega)$ or H	
Cross spectral density of forcing function and operator output			Φ_{ic}
Cross spectral density of system error and operator output			$\Phi_{\epsilon c}$
Remnant expressed as a closed loop quantity	$n_c'(t)$	$N_c'(j\omega)$ or N_c'	Φ_{nn}

* So that the various time functions, which may be stochastic, may be considered to have Fourier transforms, it should be understood that the time function used in the transform is identical to the actual time function in the intervals $-T < t < T$, and is zero elsewhere. The Fourier transform of a typical signal, $f_1(t)$, will then be defined as

$$F(j\omega) = \int_{-T}^{+T} f(t)e^{-j\omega t}dt; \quad \text{where} \quad f(t) = f_1(t); \quad -T < t < T$$
$$f(t) = 0 \text{ elsewhere.}$$

** The power spectral density of a typical signal $f_1(t)$, is given by

$$\Phi_{ff} = \lim_{T\to\infty} \frac{1}{T} F(j\omega)F(-j\omega).$$

The cross spectrum of two signals, $f_1(t)$ and $g_1(t)$ will be

$$\Phi_{fg} = \lim_{T\to\infty} \frac{1}{T} F(-j\omega)G(j\omega).$$

For more details on these quantities see, for example, 13.

Using Eqs. 1 and 4, the cross-spectral density Φ_{ic} is seen to be

$$\Phi_{ic} = \lim_{T\to\infty} \frac{1}{T}\left\{\left[\frac{Y_p I + N_c}{1 + Y_c Y_p}\right] I^*\right\}$$

$$= \left[\frac{1}{1 + Y_p Y_c}\right]\left\{ Y_p \lim_{T\to\infty} \frac{1}{T}\,[I^* I] + \lim_{T\to\infty} \frac{1}{T}\,[N_c I^*]\right\}. \quad (6)$$

The quantity, $\lim_{T\to\infty} \frac{1}{T}\,[I^* I]$ is recognized as the power spectral density of the forcing function, $\Phi_{ii}(\omega)$. The quantity $\lim_{T\to\infty} \frac{1}{T}\,[N_c I^*]$ is similarly recognized as the cross spectral density between the forcing function and the remnant, $\Phi_{in_c}(j\omega)$. This cross spectrum is zero, since by definition the remnant has no linear coherence with the forcing function. Then Eq. 6 becomes

$$\Phi_{ic} = \frac{Y_p}{1 + Y_p Y_c}\,\Phi_{ii}. \quad (7)$$

Similarly, it can be shown that the cross spectral density between the error and the input is

$$\Phi_{ie} = \frac{1}{[1 + Y_p Y_c]}\,\Phi_{ii}. \quad (8)$$

The open loop describing function Y_p can be obtained by dividing Eq. 7 by Eq. 8, and is given by

$$Y_p = \frac{\Phi_{ic}}{\Phi_{ie}}. \quad (9)$$

Equation 9 is fundamental in the measurement of the transfer characteristic of stable system elements. It is applicable to either random or periodic time functions, which must be stationary. If the system is linear and the forcing function a sine wave, the determination of Y_p reduces to a conventional frequency response measurement.

The spectral density of N_c is often an important "property" of an equivalent linear system. To find this quantity we can form the operator output power spectral density, noting, as before, that the forcing function remnant cross spectra are zero,

$$\Phi_{cc} = \lim_{T\to\infty} \frac{1}{T}\,[C^* C] = \lim_{T\to\infty} \frac{1}{T}\left\{\left[\frac{Y_p I + N_c}{1 + Y_c Y_p}\right]^*\left[\frac{Y_p I + N_c}{1 + Y_c Y_p}\right]\right\}$$

$$= \left|\frac{Y_p}{1 + Y_c Y_p}\right|^2 \lim_{T\to\infty} \frac{1}{T}\,[I^* I] + \left|\frac{1}{1 + Y_c Y_p}\right|^2 \lim_{T\to\infty} \frac{1}{T}\,[N_c^* N_c]$$

$$= \left|\frac{Y_p}{1 + Y_c Y_p}\right|^2 \Phi_{ii} + \left|\frac{1}{1 + Y_c Y_p}\right|^2 \Phi_{n_c n_c}. \quad (10)$$

In most cases it is easiest to consider the total output power in terms of that portion linearly coherent with the forcing function (the first term above) and a remainder or "closed loop" remnant. Since $Y_p/(1+Y_pY_c)$ is recognized as a closed loop describing function, H, and defining $|1/(1+Y_pY_c)|^2\,\Phi_{nn_c}$ as the closed loop remnant power spectrum Φ_{nn}, we can write Eq. 10 as

$$\Phi_{cc} = |H|^2\Phi_{ii} + \Phi_{nn}. \tag{11}$$

The possible sources and points of injection of the actual quantities composing the closed loop remnant will be discussed later. For the present, the term Φ_{nn} is selected to represent, without regard to origin, that portion of the operator's output power density which is not linearly coherent with the forcing function.

Besides the describing function and the remnant, it is desirable to have some alternate means of relating the amount of operator output power "explained" by the action of the quasi-linear describing function to the total output power. We can define such a quantity by forming the ratio of the linearly correlated output power and total output power spectra. This fraction is the square of the linear correlation, ρ, and is given by

$$\rho^2 = \frac{|H|^2\Phi_{ii}}{\Phi_{cc}} = 1 - \frac{\Phi_{nn}}{\Phi_{cc}}. \tag{12}$$

The linear correlation can also be expressed as

$$\rho = \frac{|\Phi_{ic}|}{\sqrt{\Phi_{ii}\Phi_{cc}}} \tag{13}$$

for,

$$\rho^2 = \frac{|H|^2\Phi_{ii}}{\Phi_{cc}} = \frac{\Phi_{ic}\Phi_{ic}^*}{\Phi_{ii}^2}\frac{\Phi_{ii}}{\Phi_{cc}} = \frac{|\Phi_{ic}|^2}{\Phi_{ii}\Phi_{cc}}.$$

It is apparent from Eq. 12 that for a linear system (remnant, Φ_{nn}, equal to zero) the value of ρ would be unity. In general, the value of ρ is a good measure of the importance of the remnant. As will be seen later, the linear correlation is not necessarily a good measure of linearity, although a value near unity does imply that the system is "almost" linear.

All of the quantities presented above are applicable to most general systems where servo analysis methods have a reasonable chance of success. They are essential, of course, in attempting to describe human operation because of the enormous adaptability of man as a control element.

SUMMARY OF HUMAN RESPONSE DATA FOR COMPENSATORY TASKS

Introduction

A. Tustin first noted that operators in manual control systems with random appearing forcing functions exhibit a type of behavior directly analogous to that of the equalizing elements inserted into servo systems to improve the over-all performance of the system. By its nature, this equalization appears to be a more or less continuous operation performed on data observed by the external senses. Since the operator must adapt his equalizing behavior to cope with controlled elements having widely diverse dynamic characteristics, it is reasonable to classify describing function experiments on the basis of the controlled elements which "mold" the operator's characteristics. The discussion of operator describing functions in compensatory tasks has therefore been divided into two major parts: the first will treat past experimental describing function data on compensatory systems with unity gain controlled element dynamics, that is, $Y_c = 1$, and the second will consider describing functions for operators in systems with more complex controlled element dynamic characteristics. To complete the characterization of the operator, data on the remnant are also presented, though pertinent data are sparse.

Operator Describing Functions in Compensatory Tasks with No Controlled Element Dynamics

The experimental data considered for $Y_c = 1$ were taken directly from the efforts of three investigators and their colleagues, specifically J. I. Elkind (14, 15), L. Russell (16) and E. S. Krendel (17). The primary experimental variables in these researches were the type of controller and the forcing function characteristics. The details of the forcing functions used are summarized later in Tables II and III. These experimental data have been replotted in Bode diagram form (when the original data were not in this form) and describing functions for the operator have been obtained from curve fits to these plots. The describing functions as derived from these curves are the basis for making analytical estimates from which one can determine the influence of the forcing function upon the human operator's dynamic characteristics. To the extent that the curve fitted analytical model approximates the operator, we have insight into the laws which actually govern the over-all behavior of the operator.

The most desirable types of linear models are transfer functions made up of ratios of rational polynomials, with a pure time delay term also allowed. Such transfer function forms are favored because they are simple, well understood, and adequate for approximating any of a large number of frequency response characteristics. Unfortunately, since the experimental data covered a restricted frequency range, with

poor resolution, it is possible to fit functions to the limited number of data points to an arbitrarily defined degree of accuracy. A painstaking and accurate fitting procedure is not warranted because of both the measurement and data reduction errors and the variation between subjects whose data were averaged. The appropriate model to use in fitting a set of data is, therefore, the simplest transfer function form which is reasonably consistent with all of the data trends and is appropriate for both amplitude and phase characteristics. The need for consistency between amplitude and phase forces compromises in the estimates of parameters.

As general criteria for the types of transfer function forms acceptable we list the following which are consistent with all known human operator data in compensatory tasks of various kinds.

1. The operator's describing function must approach zero at infinite frequency, that is, the transfer characteristics are fundamentally those of a low pass filter.
2. The describing function must be finite at zero frequency because of physical limitations on the force-displacement capabilities of the operator.
3. A pure time delay, $e^{-\tau s}$, must be included to account for a reaction time delay.
4. When the experimental closed loop system is stable, the approximate fitted describing function must define a stable closed loop system.

As an example of this fitting procedure, we shall consider the case of Elkind's data for forcing functions with rectangular spectra. (The final describing functions for this case are summarized in Table III.)[5] The simplest transfer function form that satisfied the foregoing requirements for these data is

$$Y_{p_1} = \frac{Ke^{-\tau_1 j\omega}}{(T_0 j\omega + 1)} \tag{14}$$

where K is the d-c. gain, τ_1 a pure time delay, and T_0 is a time constant representing a low frequency lag. Since the pure time delay $e^{-\tau_1 j\omega}$ does not affect the amplitude characteristics, the amplitude ratio curves are fitted by

$$|Y_{p_1}| = \left[\frac{T_0 j\omega + 1}{K}\right]. \tag{15}$$

Approximate values for K and T_0 can be determined from the raw data

[5] The original data upon which all of our discussion is based can be found in the original sources in various forms. They are also presented in (1) in Bode diagram form. Because of the enormous amount of data involved (Tables II and III summarize over forty separate experimental averages derived from over 300 individual runs), it is patently impossible to show even a small variety here.

amplitude ratio plots with the chief criterion the visual appearance of a good fit.

The phase associated with the term $(T_0 j\omega + 1)^{-1}$ can then be subtracted from the raw data phase to isolate the phase lag due to $e^{-\tau_1 j\omega}$. On a linear plot of phase *versus* frequency, this residual phase should be approximately linear with a slope of $-\tau_1$. When the residual phase is not approximately linear, it is necessary to select other values of T_0 (and hence K), and repeat the process until satisfactory results are attained.

While the low frequency approximation of Eq. 14 fits much of Elkind's data quite well, an analytical closed loop transfer function developed from this equation, with numbers derived from the experimental data points, would indicate an unstable condition at frequencies well beyond the measurement bandwidth. This was the case for all of Elkind's lower cutoff forcing functions. Since there was no evidence of a closed loop instability in the experimental data, the operator presumably introduces more attenuation at high frequencies than is indicated by Eq. 15. Although the experimental data do not define the high frequency portion of an approximate transfer function, it is appropriate to "stabilize" the closed loop system by adding a lag to Eq. 14 and modifying the value of τ_1 to τ, that is, changing Eq. 14 to,

$$Y_p(j\omega) \simeq \frac{Ke^{-\tau j\omega}}{(T_0 j\omega + 1)(T_1 j\omega + 1)}. \tag{16}$$

If $1/T_1$ is much greater than any of the frequencies measured, then the low frequency behavior of the transfer function of Eq. 16 can be made the same as that of Eq. 14 if $T_1 + \tau$ is approximately equal to τ_1. This addition of a simple lag, $T_1 j\omega + 1$, is not a unique way to place a stability bound upon the possible higher frequency behaviors; any combination of lags and leads leading to a stable system would be suitable if it satisfied the relationships,

$$\tau + \sum_i T_i - \sum_k T_k = \tau_1 \qquad \frac{1}{\tau_i} \gg \omega_{co}; \quad \frac{1}{T_k} \gg \omega_{co} \tag{17}$$

where the T_i are lags, the T_k are leads and ω_{co} is the forcing function cutoff frequency (or an approximation thereto) in rad/second. The single lag represented by T_1 is, however, the simplest form which can be added to Eq. 14 to assure stability and consistency with the measured low frequency response. Even with this procedure it is unfortunately not possible to find unique values of τ and T_1 from the low frequency experimental data.

The results of processing the various experimental data for systems with unity controlled element dynamics in the fashion described above, are shown in Tables II and III.[6] These tables delineate the experi-

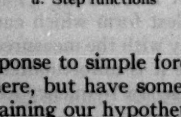
[6] The symbol s has been used in all summary tables for $j\omega$.

mental conditions, the frequency ranges of measurement, the average linear correlations, and the approximate describing function fits to the data. Table I is also included to provide a summary of operator re-

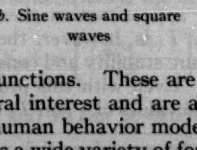
TABLE 1.—*Summary of Operator Describing Functions in Compensatory Tasks. Simple Tracker, $Y_e = 1$, Simple Forcing Functions.*

General Control Task	"Best Fit" Human Operator Transfer Function	Investigators and Remarks
Simple following with pencil, stick, or wheel (forcing function *a*)	Closed loop transfer function $\frac{e^{-\tau s}}{\frac{T^2s^2}{K}+\frac{2\zeta Ts}{K}+\frac{s}{K}+1} \doteq \frac{e^{-\tau s}}{T_N s+1}$	Goodyear (**35**) also: Searle & Taylor
	$\zeta \doteq 0.5 \quad K \doteq 6.8\text{–}7.8 \text{ sec}^{-1}$. $\tau \doteq 0.25$ sec. $T = 0.042$ sec. $T_N \doteq 1/K$	Cheatham Mayne Ellson & Hill
Simple following with pencil or wheel (forcing function *b*)	"*Nonsynchronous*" response, closed loop $\frac{1}{\frac{T^2s^2}{K}+\frac{2\zeta Ts}{K}+\frac{s}{K}+1}$ with constants as above	Goodyear (**19, 35**) Mayne, Ellson & Gray, Cheatham
	"*Synchronous*" response, closed loop $K(\omega)$	
Simple following with pencil, stick, or wheel (forcing function *c*)	Same as for single steps Though τ can approach 0.20 sec. as steps get closer together and appear more random	Goodyear (**19, 35**) Mayne small range effect; Searle & Taylor Ellson & Wheeler Slack

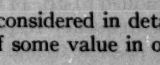
Types of forcing functions:



a. Step functions



b. Sine waves and square waves



c. Sequence of steps

sponse to simple forcing functions. These are not considered in detail here, but have some general interest and are also of some value in obtaining our hypothetical human behavior models.

Tables II and III shows a wide variety of forcing function conditions represented by their power spectral densities, Φ_{ii}. The broadest coverage is due to Elkind (Table III), who had remarkable success establishing relationships between the fitted describing function param-

TABLE II.—*Summary of Operator Describing Functions in Compensatory Tasks. Simple Tracker, $Y_c = 1$, Random Appearing Forcing Functions.*

General Control Task	"Best Fit" Human Operator Transfer Function	Frequency Range of Human Operator Measurements	Average Linear Correlation	Investigators and Remarks
Simple tracker with spring restrained aircraft control stick in aircraft cockpit mockup (forcing function *d*)	$\frac{Ke^{-\tau s}(T_L s+1)}{(T_I s+1)(T_N s+1)}$ Corner Freq. $1/T_I$ $1/T_N$ $1/T_L$ τ K 1 .04 1.5 0.5 0.15 100 2 .11 4.55 2.0 0.20 40 4 0.2 11.0 3.0 0.25 1.5	0.4 to 4.0 rad/sec.	0.7 to 0.8	Franklin (17, 1), these data taken for "lateral" control; "longitudinal" control was exercised simultaneously
Simple tracker handwheel type control with no restraints (forcing function *e*)	$\frac{Ke^{-\tau s}(T_L s+1)}{(T_I s+1)(T_N s+1)}$ $\tau \doteq 0.3$; T_N effect included in τ Forcing Function $1/T_I$ $1/T_L$ K Low speed .103 13.3 52.5 Medium speed .37 5.0 11.0 High speed .62 3.7 2.0	Same frequencies as shown in forcing function	0.9	Russell
Simple tracker aircraft type handwheel control, with spring restraint, operated in pitch (push-pull) direction (forcing function *f*)	$\frac{Ke^{-\tau s}(T_L s+1)}{(T_I s+1)}$ $\tau \doteq 0.2$ Controlled Element Gain $1/T_I$ $1/T_L$ K 5 .027 2.7 10 10 .027 2.7 7	0.5 to 3 rad/sec.	$K_c=10$: 0.98 to 0.6 $K_c=5$: 0.88 to 0.45	Hall, lateral axis airplane dynamics also being controlled. Average decreased with frequency

Types of forcing functions:

d. White noise through third order binomial filter giving available corner frequencies of 1, 2, and 4 rad/sec.

.277 1.21 (Low)
.66 2.87 (High)
1.38 6.03 (Medium)
.741 1.80 (Low)
1.68 4.27 (Med.)
3.54 8.98 (High)

e. Superposition of four sinusoids, rad /sec

f. White noise through third order binomial filter, corner at 1 rad/sec

eters and other parameters defining the forcing function characteristics. These are also given in Table III. Some work has been done (1) in an endeavor to find underlying criteria which may lead to these relationships. While these attempts were not completely successful, one can show that the conventional servo synthesis method of parameter adjustment to minimize the rms error (13) leads to analogous results.

TABLE III.—*Summary of Operator Describing Functions in Compensatory Tasks Showing Effects of Forcing Functions. Simple Tracker, $Y_c = 1$, Random Appearing Forcing Function. These Data Taken from Elkind, in a Control Task Consisting of Simple Following with a Pip Trapper.*

Type of Forcing Function	"Best Fit" Human Operator Transfer Function	Frequency Range of Human Operator Measurements	Average Linear Correlation, ρ_a
Function made up of 40–120 sinusoids, giving any desirable rectangular spectra 1″ rms amplitude (see *g* below)	$\frac{Ke^{-\tau s}}{(T_I s+1)(T_N s+1)}$ where: $K \doteq 2.2/f_{co}^2$; $K \doteq 9.42 T_I$ *Actual fitted data:*	Highest frequency equal to forcing function bandwidth, lowest frequencies given below in rad/sec.	R.16–.995 R.24–.99 R.40–.995 R.64–.98 R.96–.92 R1.6–.75 R2.4–.58
Approximate white noise through 1st, 2nd, and 3rd order binomial filters as shown below. Highest frequency in forcing functions was 2.88 cps. 1″ rms amplitude (see *h* below)	$\frac{Ke^{-\tau s}}{(T_I s+1)}$ where: $K \doteq \frac{3.8}{\sigma_f e^{7/0.33}}$, or $K \doteq \frac{0.39}{\sigma_f f}$; $K \doteq 9.42 T_I$ $\sigma_f = \left[\frac{\int_0^\infty f^2 \Phi_{ii} df}{\int_0^\infty \Phi_{ii} df} - (f)^2\right]^{\frac{1}{2}}$; $f = \frac{\int_0^\infty f \Phi_{ii} df}{\int_0^\infty \Phi_{ii} df}$ *Actual fitted data:*	Range (rad/sec) Filter 1st order .25 to 18 2nd order .25 to 12 3rd order .25 to 6	Filter ρ_a 1st order 0.9 2nd order 0.96 3rd order 0.97

Forcing Function	K(db)	τ	$1/T_I$	$1/T_N$
R.16	34.5	.110	.22	1.885
R.24	31.5	.104	.314	6.22
R.40	22.5	.133	.785	12.3
R.64	15.0	.150	1.73	30.3
R.96	6.5	.139	3.65	∞
R1.6	−0.6	.122	3.77	∞
R2.4	−3.0	.116	1,885	∞

Forcing Function and Filter	K(db)	$1/T_I$	τ
F1; 1st order	10.5	1.13	.139
F2; 2nd order	25.0	0.314	.126
F3; 3rd order	33.0	.1885	.178

TABLE III.—*Continued.*

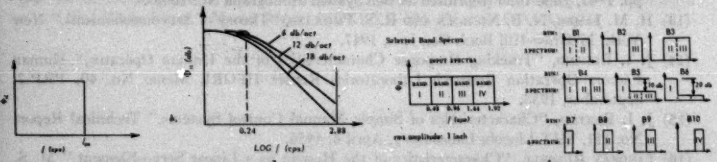
Type of Forcing Function	"Best Fit" Human Operator Transfer Function	Frequency Range of Human Operator Measurements	Average Linear Correlation, ρ_a
24 to 144 sinusoidal components giving forcing function spectra shown (No. of components shown with particular spectra in parentheses.) 1" rms (see *i* below)	$\frac{Ke^{-\tau s}}{(T_I s+1)(T_N s+1)}$ {where $Y(s)$ was translated to the origin for B3, B8, B9, B10	B1 .6 to 6 rad/sec. B2 .6 to 3 and 6.3 to 9 B3 3 to 9 B4 .6 to 8.7 B5 .6 to 9 B6 .6 to 9 B7 .6 to 3 B8 3 to 6 B9 6 to 9 B10 9 to 12	B1 .97 B2 .91 B3 .92 B4 .90 B5 .92 B6 .90 B7 .99 B8 .97 B9 .91 B10 .70

where:

$$K \doteq 9.42 T_I; \quad K \doteq \frac{3.8}{\sigma_{fc} f_{co}^{0.35}}$$

Actual fitted data:

Forcing Function	K(db)	$1/T_I$	$1/T_N$	τ
B1	9.0	4.78	∞	.153
B2	1.5	5.03	∞	.107
B3	−1.0	12.6	∞	.278
B4	−0.3	12.6	∞	.150
B5	11.1	1.88	∞	.128
B6	17.7	1.00	∞	.149
B7	23.2	.88	17.8	.100
B8	9.0	3.14	6.28	.219
B9	0.4	12.6	∞	.390
B10	−11.8	2.82	∞	1.14*



f_{co} = .16, .24, .40, .64, .96, 1.6, 2.4 cps

or

$\omega_{co} \doteq$ 1, 1.5, 2.5, 4, 6, 10, 15 rad/sec.

(*To be continued*)

REFERENCES†

(1) D. T. McRuer and E. S. Krendel, "Dynamic Response of Human Operators," WADC TR 56-524, August, 1957.
(2) "Handbook of Human Engineering Data," Tufts College Institute for Applied Experimental Psychology, Human Engineering Report SDC 119-1 (Nav Exos P-643), second edition, December, 1951.
(3) R. G. E. Epple, "The Human Pilot," Northrop Aircraft, Inc., BuAer Report AE-61-4III, August, 1954.
(4) J. W. Senders and L. J. Fogel, "The Human as a Dual Mode Relay Servomechanism," Man-Machine Servo Conference, Franklin Institute, Feb. 14-15, 1955.
(5) H. L. Platzer and E. S. Krendel, "Nonlinear Techniques for Studying Human Operator Performance," Man-Machine Servo Conference, Franklin Institute, Feb. 14-15, 1955.
(6) H. L. Platzer, "The Phase-Plane as a Tool for the Study of Human Behavior in Tracking Problems," WADC TR 55-444, November, 1955.
(7) H. L. Platzer, "A Nonlinear Approach to Human Tracking," Franklin Institute Report I-2490-1, Contract Nav 1571(00), December, 1955.
(8) J. R. Moore, "Combination Open-Cycle, Closed-Cycle System," *Proc. IRE*, Vol. 39, pp. 1421-1432 (1951).
(9) N. Kryloff and N. Bogoliuboff, "Introduction to Nonlinear Mechanics," translated by S. Lepschetz, Princeton University Press, 1943.
(10) R. C. Booton, Jr., "The Analysis of Nonlinear Control Systems with Random Inputs," Proc. Symp. Nonlinear Circuit Analysis, Brooklyn Polytechnic Institute, April, 1953.
(11) R. C. Booton, Jr., "Nonlinear Control Systems with Random Inputs," *Trans. IRE PGCT*, Vol. CT-1, pp. 9-17 (1954).
(12) H. Press and J. W. Tukey, "Power Spectral Methods of Analysis and Their Application to Problems in Airplane Dynamics," Part IV-C, Vol. IV, AGARD Flight Test Manual. pp. 1-41, June 1956 (reprinted as Bell System Monograph No. 2606).
(13) H. M. James, N. B. Nichols and R. S. Phillips, "Theory of Servomechanisms," New York, McGraw-Hill Book Co., Inc., 1947.
(14) J. I. Elkind, "Tracking Response Characteristics of the Human Operator," Human Factors Operation Research Laboratories Report HFORL Memo No. 40, PRP-2, September, 1953.
(15) J. I. Elkind, "Characteristics of Simple Manual Control Systems," Technical Report No. 111, MIT Lincoln Laboratory, April 6, 1956.
(16) Lindsay Russell, "Characteristics of the Human as a Linear Servo-Element," M. S. thesis, MIT, May 18, 1951.
(17) E. S. Krendel and G. H. Barnes, "Interim Report on Human Frequency Response Studies," WADC TR 54-370, June, 1954.
(18) A. Tustin, "The Nature of the Operator's Response in Manual Control and Its Implications for Controller Design," *J. Inst. Elec. Engrs.*, Vol. 94, Part II A, pp. 190-202 (1947).
(19) "Final Report: Human Dynamic Study," Goodyear Aircraft Corp., Report No. GER-4750, April 8, 1952.
(20) "Investigation of Vestibular and Body Reactions to the Dynamic Response of the Human Operator," Final Report, BuAer Report AE-61-6 and Goodyear Aircraft Corp., Report GER 5452, November 25, 1953.
(21) "Investigation of Control 'Feel' Effects on the Dynamics of a Piloted Aircraft System," Goodyear Aircraft Corp. Report GER 6726, April 25, 1955.
(22) N. D. Diamantides and A. J. Cacioppo, "Human Response Dynamics; Geda Computer Application," Goodyear Aircraft Corp. Report GER 8033, January 8, 1957.
(23) I. A. M. Hall, "Effect of Controlled Element on the Human Pilot," WADC TR 57-509, October, 1957.

† Editor's Note: The majority of the above reports are available to qualified users of ASTIA, Arlington Hall Station, Arlington 12, Va.

(24) H. P. BIRMINGHAM AND F. V. TAYLOR, "A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems," NRL Report 43333, April 7, 1954.
(25) H. P. BIRMINGHAM AND F. V. TAYLOR, "A Design Philosophy for Man-Machine Control Systems," *Proc. IRE*, Vol. 42, pp. 1748–1758 (1954).
(26) H. P. BIRMINGHAM, A. KAHN AND F. V. TAYLOR, "A Demonstration of the Effects of Quickening in Multiple-Coordinate Control Tasks," NRL Report 4380, June 23, 1954.
(27) J. G. HOLLAND AND J. B. HENSON, "Transfer of Training Between Quickened and Unquickened Tracking Systems," NRL Report 4703, February 3, 1956.
(28) W. D. GARVEY AND L. L. MITNICK, "An Analysis of Tracking Behavior in Terms of Lead-Lag Errors," NRL Report 4707, February 16, 1956.
(29) L. V. SEARLE AND F. V. TAYLOR, "Studies of Tracking Behavior: I. Rate and Time Characteristics of Simple Corrective Movements," *J. Exptl. Psych.*, Vol. 38, pp. 613–631 (1948).
(30) D. G. ELLSON AND H. HILL, "The Interaction of Response to Step Function Stimuli: I. Opposed Steps of Constant Amplitude," AMC Memo Report No. MCREXD-694–2P, November 19, 1948.
(31) R. MAYNE, "Some Engineering Aspects of the Mechanism of Body Control," *Elec. Eng.*, Vol. 70, pp. 207–212 (1951).
(32) D. C. CHEATHAM, "A Study of the Characteristics of Human Pilot Control Response to Simulated Aircraft Lateral Motions," NACA Report 1197, 1954.
(33) D. G. ELLSON AND L. WHEELER, "The Range Effect," Air Force Technical Report 5813, May, 1949.
(34) C. W. SLACK, "Some Characteristics of the 'Range Effect'," *J. Exptl. Psych.*, Vol. 46, pp. 76–80 (1953).
(35) "A Proposal to Study the Dynamic Characteristics of Man as Related to the Man-Aircraft System," Goodyear Aircraft Corp. Report GER-3006-A, May 2, 1950.
(36) D. G. ELLSON AND F. GRAY, "Frequency Responses of Human Operators Following a Sine Wave Input," Memo Report MCREXD-694–2N, USAF AMC, 1948.

Simple Electronic Circuit Simulates Living Nerve Cell.—A simple electronic circuit that simulates some functions of the individual biological nerve cell, or neuron, has been developed at Bell Telephone Laboratories. Numbers of these artificial cells are being combined into experimental networks that are roughly analogous to the nerve systems of the eye and ear.

Bell Labs scientists are especially interested in discovering how visual and auditory nerves function and how their signals are interpreted by the brain. Knowledge derived from such basic research may lead, as it has in the past, to better and more economical communication.

The gross function of the nerve cell —transmission of electrical pulses in response to stimuli, and only to those stimuli that meet certain conditions— has been simulated by a simple circuit. L. D. Harmon, who initiated the project, has combined the cells into groups that simulate simple functions of the eye, and similar experiments with ear models are being started by W. A. Van Bergeijk.

The neuron circuit, designed by R. M. Wolfe, fires electrical pulses of standard amplitude and duration, just as a biological cell usually does. If the circuit is driven by a constant stimulus, simulating receptor cells as in the eye or ear, trains of pulses are emitted. A higher intensity of excitation increases the frequency of pulsing; and when the neuron is excited continuously, the frequency of pulses can be made to decrease with time, exhibiting accommodation as a living nerve cell does.

Input excitation must, as in a biological cell, surpass a threshold value, and the cell will integrate two or more input pulses below threshold value to cause firing. A particular input connection can also, while energized, inhibit firing of the neuron by other inputs. Similarly, immediately after firing, the electronic neuron's threshold rises to infinity and for a few milliseconds no input signal can fire the neuron again.

The circuit includes four transistors, thirteen resistors, and two capacitors, mounted on a three-by-four-inch, printed-circuit card for ease of handling. The pulse length it delivers, about six milliseconds, is considerably longer than that of a biological nerve, but can be shortened if desired. The cell has an integrating time-constant of two milliseconds and a refractory time-constant of about ten milliseconds, approximating time-constants of the biological neuron. Because the electronic inputs and outputs are compatible, the cells can be assembled into chains and networks.

Electronic neurons can be combined with photo-resistive cells to simulate simple functions of nerves in the retina. Some receptors, known as "on" receptors, fire only when the light intensity they receive is increasing; "off" receptors fire only when the light is decreasing; and "during" receptors fire while they receive a steady light. Flicker-fusion phenomena have also been produced. In the human eye, these can cause a sequence of flashes to be seen as continuous illumination; this property of vision is exploited in motion pictures and TV.

The mutual inhibition of cells in an array has been demonstrated experimentally. Some animals have been observed to possess this arrangement, in which a cell receiving a greater light intensity inhibits the firing of nearby cells that receive less light. This results in local sharpening of image boundary detail.

It is hoped that further imitation of nerve functions, so far as they are known, and other experiments may lead to better understanding and prediction of neurological behavior.

A UNIFIED APPROACH TO THE THEORY OF SAMPLING SYSTEMS*

BY

R. E. KALMAN[1] AND J. E. BERTRAM[2]

ABSTRACT

This paper presents a new method for the analysis of sampling (or sampled-data) systems. Unlike methods currently in use, it is not necessary to assume that the sampling operations are synchronous, performed at constant rate, and representable by means of an impulse modulator. Several different types of sampling operations are considered in detail, with the analysis proceeding similarly in each case. The paper concludes with a brief study of the stability of sampling systems and a generalization of Floquet's theorem.

INTRODUCTION

A dynamic system in which one or more variables are allowed to to change only at discrete instants of time (called *sampling instants*) is known as a *sampled-data* or more simply as a *sampling* system (1).[3] Sampling systems may be regarded as a subclass of nonstationary (time-varying) dynamic systems. Such systems arise naturally when a digital computer, a pulse or counting circuit, or any other discontinuously acting element is incorporated in a system.

Conventional Sampling

At the present time, most of the literature (1–5) dealing with the analysis and design of sampling systems is based on the following simplifying assumptions:

(*a*) All sampling operations can be adequately described by introducing a fictitious idealized component, called an *impulse modulator* (2) into the system.

(*b*) All sampling operations are performed synchronously.

(*c*) The time interval between successive samples is fixed.

In addition, it is almost always assumed that the system is linear.

A sampling system whose operation (within the accuracies demanded by engineering considerations) satisfies these simplifying assumptions will be referred to in this paper as a *conventional sampling system.*

The analysis and synthesis of conventional sampling systems have

* This work was done while both authors were in the Department of Electrical Engineering at Columbia University. The work was supported in part by the IBM Research Laboratories.

[1] RIAS, Baltimore, Md.

[2] IBM Research Laboratory, Ossining, N. Y.

[3] The boldface numbers in parentheses refer to the references appended to this paper.

been well developed and standardized in recent years with the aid of the z-transform method (1, 3–5).

Recently there has been increased interest in systems in which the sampling operation does not satisfy one or more of Assumptions (a), (b), (c). Examples of such systems are:

Nonsynchronous Sampling

All sampling operations are repeated at the same rate but occur at different instants of time. Such problems can be analyzed with the aid of the "modified" z-transform method (1, 4, 5) due to Barker. In all except the simplest cases, however, the algebraic details of the analysis become exceedingly laborious.

Multiple-Order Sampling

All sampling operations in the system are performed synchronously and a group of sampling operations is repeated every T seconds, the intervals between successive samples in the group being unequal. For example, *m^{th}-order sampling* means that once very T seconds samples are taken at times

$$t = kT, \quad kT + \tau_1, \cdots, \quad kT + \tau_{m-1}, \quad (k+1)T, \cdots \quad (k = 0, 1, \cdots)$$

where the τ_i are arbitrary constants subject to the condition $0 < \tau_1 < \tau_2 < \cdots < \tau_{m-1} < T$. Such sampling operations may arise when a digital computer is time-shared by several feedback control systems. Systems with multiple-order sampling have not been studied previously as far as the authors are aware. (Added in proof: See 20, 21.)

Multi-Rate Sampling

There may be several sampling operations in the system with fixed but unequal sampling intervals. Kranc (6) has made an extensive study of multi-rate systems in which the ratios of the various sampling periods are rational numbers. His analysis is based on a special transform method.

Noninstantaneous Sampling

Here the sampling interval is too long with respect to the time-scale of the dynamic components of the system to be adequately described by impulse modulation. An approximate solution to this problem has been obtained by Kranc (7), using transform methods. An exact solution has been published by Farmanfarma (8). See also the writers' discussion of Farmanfarma's paper.

Random Sampling

In some cases, the intervals between successive samples may be thought of as being selected according to some random scheme. Ap-

proximately random sampling may be caused by the desire to reduce susceptibility of the system to jamming or similar interference; by random time delays in the sampling operation depending on a physical measurement; by a random time delay which must elapse before a digital computer again becomes available to carry out control computations required by a particular feedback loop of a multi-loop system; and finally by inaccuracies in equipment which is designed to sample at constant intervals.

A comprehensive study of systems with random sampling has been carried out by one of the writers using mainly the methods of this paper; these results are discussed elsewhere (18). (Added in proof: See 22.)

Evidently, all five types of sampling mentioned above may be encountered in a single system.

The New Method

The objective of this paper is to present a uniform method of analysis which can be applied in a straightforward fashion to sampling systems of the types mentioned. The same method has been used recently by the writers in connection with:

(i) design of conventional sampling systems for "deadbeat" response (9);
(ii) design of an optimal relay servo operating with conventional sampling (10);
(iii) design of conventional sampling systems to be optimal in the Wiener sense (11);
(iv) systems with random sampling (18);
(v) analysis of effect of round-off errors on the dynamic performance of sampling systems incorporating a digital element (12);
(vi) analysis of conventional sampling systems; and
(vii) design of Wiener filters operating on sampled data (19).

Transform theory as used in the past to study conventional sampling systems has proved to be either unwieldy or has required elaborate modifications before it could be applied to systems of the type mentioned. By using the concepts of *state* and *transition matrix* (see the next two sections) it is possible to handle systems of the general type in a clear and uniform way. The new method yields simplifications even in the analysis and synthesis of conventional sampling systems. The method automatically eliminates one of the chief difficulties of the transform method, namely that with the latter it is difficult or cumbersome to obtain information about the behavior of the system at any time other than the sampling instants. The method is in accordance with modern trends in engineering science; it devotes paper-and-pencil studies only to the analytic aspects of system problems, leaving the drudgery of numerical computation to be performed by a digital computer.

Finally, the formulation of system equations according to the point of view of this paper is a natural and necessary step preliminary to the study of nonlinear sampling systems.

ELEMENTS OF LINEAR SAMPLING SYSTEMS

Because of the generality of the method of analysis to be presented, careful attention must be paid to the definition of the individual elements of the system. The elements of a sampling system to which the method applies are of the following type:

Dynamic Elements

These are elements whose present output depends not only on the present value of the input but also on past values of the input. We will be concerned with two particular classes of dynamic elements:

Linear Continuous Dynamic Elements. These are assumed to be described by ordinary linear differential equations with constant coefficients. They represent, in an idealized fashion, such diverse physical entities as motors, ships, airplanes, missiles, chemical reactors, electrical networks, etc.

Linear Discrete Dynamic Elements. These are governed by linear difference equations with constant coefficients. Such elements are peculiar to sampling systems. They are an idealization of linear numerical computations performed by digital computers incorporated in the system. (Networks of delay lines used, for instance, to process radar data can be interpreted similarly.) A digital computer cannot operate directly on a continuous signal, but only on sequences of numbers, that is, on a *discrete signal.* Thus discrete dynamic elements inherently involve the operation of *sampling*, namely the conversion of continuous signals into discrete signals.

Sample-and-Hold Elements

These represent, again in an idealized fashion, the operation of "sampling" various types of information (measuring a radar return pulse, sampling the values of any continuous signal, etc.) and then applying the sampled information through a "smoothing" filter to the input of some continuous dynamic element. A sample-and-hold element is necessary whenever the output of a discrete dynamic element is to be applied to the input of a continuous dynamic element, since the output of a discrete dynamic element can only be observed through the process of sampling it (that is, instructing the digital computer to "write" out numbers at specific instants of time) and the input of a continuous dynamic element cannot be a discrete signal. Sometimes a sample-and-hold element is inserted in the system because, paradoxically, sampling may improve the dynamic performance of the system (1, 9).

The simplest "smoothing" operation consists simply of holding the last sampled value until the next sample is available. If the input to the sample-and-hold element is $u(t)$, its output $v(t)$ is given by

$$v(t_k + \tau) = u(t_k), \quad 0 < \tau \leqq t_{k+1} - t_k \tag{2}$$

where t_k ($k = 0, 1, 2, \cdots$) is the k^{th} *sampling instant* and τ is the time elapsed since the last sample was taken.

The manner in which the sampling instants are to be selected will be discussed in detail in the next section.

Imperfect Hold. In some cases, it is difficult to maintain the output of the sample-and-hold element exactly at the sampled value as required by Eq. 2. Instead of (2), the input-output relations of the sample-and-hold element may take the form

$$v(t_k + \tau) = (\exp \tau/T_s)u(t_k), \quad 0 < \tau \leqq t_{k+1} - t_k \tag{2-A}$$

where T_s is the time-constant of the sample-and-hold element. Of course T_s should be as large as possible compared to the average time interval between successive samples.

The modification represented by (2-*A*) can be easily incorporated in the general analysis which follows and therefore need not be discussed further.

Noninstantaneous Sampling. In this case, the sample-and-hold element transmits the input signal without change for U seconds following t_k; thereafter, the output is held constant at the last sampled value of the input. Therefore in this case we have

$$\begin{aligned} v(t_k + \tau) &= u(t), && 0 < \tau \leqq U \\ &= u(t_k + U), && U \leqq \tau \leqq t_{k+1} - t_k. \end{aligned} \tag{2-B}$$

Better Smoothing. Sometimes the smoothing provided by the sample-and-hold element is inadequate. In such cases, additional smoothing may be provided by cascading a linear dynamic element (say, n cascaded integrators) after the sample-and-hold element. This modification obviously requires no special consideration.

Interconnecting Elements

These represent idealized operations on the outputs of the dynamic elements, such as multiplication by a constant, differentiation, as well as linear combinations of these operations. Differentiation can occur only in connection with the output of continuous dynamic elements. See in this connection, *Remark* (*a*, ii), p. 413.

The preceding classification of elements is valid also in the case of nonlinear sampling systems. The latter are characterized by the fact that the interconnecting elements have nonlinear input-output relations; in other words, the principle of superposition is not valid.

To illustrate the manner in which the above elements are interconnected, consider the hypothetical feedback system shown in Fig. 1. The system consists of two continuous dynamic elements (*CDE*), one discrete dynamic element (*DDE*) and two sample-and-hold elements (*SHE*). The boxes with the symbols K, d/dt, and Σ denote interconnecting elements corresponding to multiplication by a constant, differentiation, and summation. The rest of the interconnections are indicated simply by the lines connecting the various boxes. The quantities x_7 (the output of the discrete dynamic element) and x_8, x_9 are discrete signals. All other signals are continuous. The *system input* is denoted by $r(t)$ and the *system output* by $x_1(t)$.

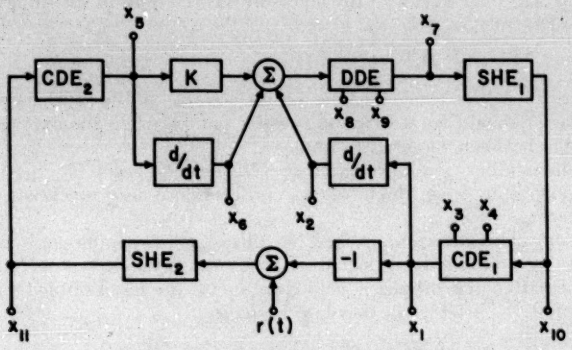


FIG. 1. A typical sampling system.

DESCRIPTION OF THE SAMPLING OPERATION

As mentioned in the preceding section, sampling operations occur in discrete dynamic elements as well as in sample-and-hold elements. Let us consider a single sampling operation. The k^{th} *sampling instant* is denoted by t_k $(k = 0, 1, \cdots)$. The interval between successive samples

$$T_k = t_{k+1} - t_k \tag{3}$$

is called the k^{th} *sampling period*. The various sampling operations may be characterized as follows:

Conventional Sampling. T_k = const. for all k.

Nonsynchronous Sampling. If t_k, t_k' denote the instants at which two different sampling operations occur, then $t_k' = t_k + U$ for all k, where U is some positive constant.

Multiple-Order Sampling. The sampling period is a periodic function of k; in other words, $T_k = T_{k+q}$ where q is a positive integer.

Multi-Rate Sampling. If T_k, T_k' denote the sampling periods of

two different sampling operations, T_k = const. and T_k' = const. for all k, but $T_k \neq T_k'$.

Noninstantaneous Sampling. This can occur only in connection with the sampling operations of the sample-and-hold elements, whose input-output relations in this case are given by (2-*B*).

Random Sampling. Both t_k and T_k are random variables.

The various types of sampling operations are depicted schematically in Fig. 2.

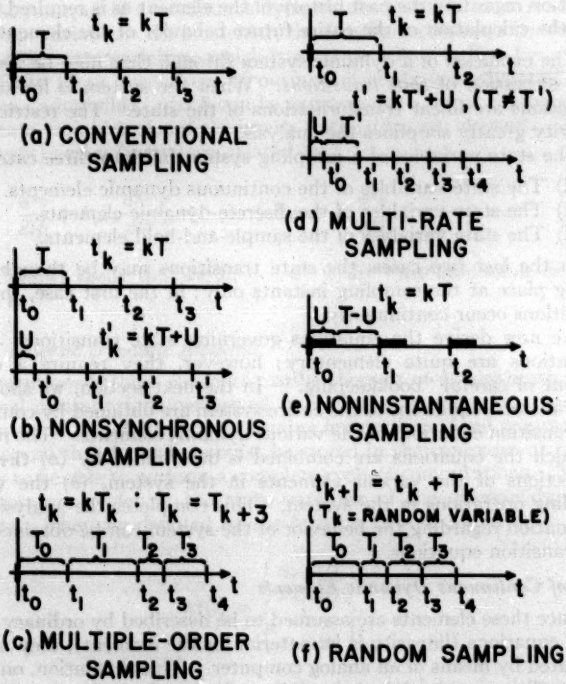


FIG. 2. The sampling operation.

Notation Convention. The following convention is used to denote the change in various quantities as a result of the sampling operations:

If $x(t)$ is some quantity whose value changes due to a sampling operation, and if t_k is the sampling instant, then we denote by $x(t_k)$ the value of $x(t)$ just before and by $x(t_k^+)$ the value of $x(t)$ just after the sampling operation. In other words, the sampling operation is thought of as taking place in an interval $t_k < t < t_k + \epsilon$, where ϵ is an arbitrarily

small positive number. Note that this convention is in strict accordance with the inequalities used in (2), (2-*A*), and (2-*B*).

MATHEMATICAL DESCRIPTION OF DYNAMIC ELEMENTS

The mathematical description and analytic study of a dynamic system is greatly facilitated by focusing attention on the concept of *state*.

> *Intuitive Definition*: The *state* of a dynamic element is a set of numbers (called *state variables*) which contain as much information regarding the past history of the element as is required for the calculation of the entire future behavior of the element.

The evolution of a dynamic system through time may be visualized as a succession of *state transitions*. When the system is linear, these transitions are linear transformations of the state. The restriction to linearity greatly simplifies the analysis.

The state variables of a sampling system fall into three categories:

(i) The state variables of the continuous dynamic elements.
(ii) The state variables of the discrete dynamic elements.
(iii) The state variables of the sample-and-hold elements.

In the last two cases, the state transitions may be thought of as taking place at the sampling instants only; in the first case, the state transitions occur continuously.

V.2
195

We now derive the equations governing state transitions. These derivations are quite elementary; however, they require a certain amount of careful "bookkeeping." In the next section, we show how the transition equations of the entire system are obtained by combining the transition equations of the various dynamic elements. The manner in which the equations are combined is determined by (*a*) the interconnections of the various elements in the system, (*b*) the various sampling operations in the system. This completes the analysis; any information regarding the behavior of the system can be obtained from the transition equations.

State of Continuous Dynamic Elements

Since these elements are assumed to be described by ordinary differential equations (linearity is immaterial for the moment), they may be simulated by means of an analog computer. This simulation, on paper or in reality, may be done in a variety of ways; in each case, exactly γ integrators will be required, γ being the sum of the orders of the differential equations governing the various continuous dynamic elements. The outputs of the γ integrators at time t are denoted by the γ-vector ($\gamma \times 1$ matrix),

$$\mathbf{x}^c(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_\gamma(t) \end{bmatrix}. \tag{4}$$

In the absence of inputs, the signals generated in the analog computer after $t = t_0$ depend only on the voltages stored on the integrators at $t = t_0$ (that is, on $\mathbf{x}^c(t_0)$); thus $\mathbf{x}^c(t)$ is clearly a correct definition of the state of the continuous dynamic elements (13).

In Fig. 1, the state variables of the first continuous dynamic element are denoted by $x_1, \cdots, x_4$; those of the second by x_5, x_6. Thus $\gamma = 6$.

Remark a. It can be easily shown from the elementary theory of linear differential equations that

(i) Any output of a continuous dynamic element is a linear function of the state variables. In particular, the output may always be chosen as one of the state variables, which is what was done in Fig. 1.

(ii) Only those derivatives of outputs of continuous dynamic elements are "admissible" which can be expressed as a linear combination of the state variables of and the inputs to the dynamic element. Thus if v (output) and u (input) are related by

$$\frac{d^p v}{dtp} + b_{p-1}\frac{d^{p-1}v}{dt^{p-1}}, + \cdots + b_0 v = a_q \frac{d^q u}{dt^q} + \cdots + a_0 u$$

then it follows easily that $dv/dt, \cdots, d^{p-q}v/dt^{p-q}$ are admissible derivatives of v.

State of Discrete Dynamic Elements

As mentioned, these elements can be thought of as idealizations of numerical computations performed by digital computers incorporated in the system. Excluding programming instructions, fixed parameters, etc., suppose that δ numbers must be stored in the computer at time t in order to be able to perform a complete cycle of computations. These δ numbers are evidently the state variables of the discrete dynamic elements. They are denoted by the δ-vector ($\delta \times 1$ matrix)

$$\mathbf{x}^d(t) = \begin{bmatrix} x_{\gamma+1}(t) \\ x_{\gamma+2}(t) \\ \vdots \\ x_{\gamma+\delta}(t) \end{bmatrix}. \tag{5}$$

In Fig. 1, the state variables of the discrete dynamic element are denoted by x_7, x_8, x_9. Thus $\delta = 3$.

State of Sample-and-Hold Elements

From (2) we see that the state of a sample-and-hold element is identical with its output. Assuming there are σ sample-and-hold elements in a given system, we denote these state variables by the σ-vector

($\sigma \times 1$ matrix)

$$\mathbf{x}^s(t) = \begin{bmatrix} x_{\gamma+\delta+1}(t) \\ x_{\gamma+\delta+2}(t) \\ \vdots \\ x_{\gamma+\delta+\sigma}(t) \end{bmatrix} \tag{6}$$

which represents the state of the entire sampling system.

In Fig. 1, the state variables of the sample-and-hold elements are x_{10}, x_{11}. Thus $\sigma = 2$.

State of Entire System

By combining the state variables defined by (4), (5), and (6), we obtain the n-vector ($n \times 1$ matrix)

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} \mathbf{x}^c(t) \\ \mathbf{x}^d(t) \\ \mathbf{x}^s(t) \end{bmatrix}$$

which represents the state of the entire sampling system. The integer $n = \gamma + \delta + \sigma$ is called the *order* of the sampling system.

Transition Equations of Sample-and-Hold Elements

The state transitions take place (discontinuously) only at sampling points (see (2)):

$$x_i(t_k^+) = \sum_{j=1}^{\gamma+\delta} s_{ij}x_j(t_k) + s_{i,n+1}r(t_k), \quad (i = \gamma + \delta + 1, \cdots, n) \tag{7-i}$$

where the s_{ij} are constants; t_k refers to the sampling instant of the i^{th} sample-and-hold element (which may or may not coincide with the sampling instants of the other elements). Equations 7 express the fact that, in general, the input to a sample-and-hold element is a linear combination of the state variables of the continuous and discrete dynamic elements as well as the system input $r(t)$.

Transition Equations of Discrete Dynamic Elements

Here

$$x_i(t_k + \tau) = \sum_{j=1}^{\gamma+\delta} d_{ij}x_j(t_k) + d_{i,n+1}r(t_k),$$

$$\tau \geqq \tau_i \ (i = \gamma + 1, \cdots, \gamma + \delta) \tag{8-i}$$

where the d_{ij} are constants and τ_i is the time required for the computations indicated by (8) to be completed by a digital computer; t_k is the instant of time at which the variable quantities entering on the right-hand side of (8) have been sampled (t_k may be the same for several of the Eqs. 8-i).

If the results of the computations indicated by (8-i) are not to be called for (as must be the case in a well-designed system) until more than τ_i seconds after the sampling instant, t_k, then we may just as well assume that the state transitions take place instantaneously at t_k, in other words, the left-hand side of (8-i) may be denoted by $x_i(t_k^+)$ (see "Notation Convention," p. 411). Often the computing time τ_i is very short compared to the system dynamics and may be disregarded altogether.

Transition Equations of Continuous Dynamic Elements

Let us define, in conjunction with the analog computer setup introduced in a previous section (see p. 412), the constants

a_{ij} = coefficients with which the output of the j^{th} integrator is fed back to the input of the i^{th} integrator. (See Fig. 3.)

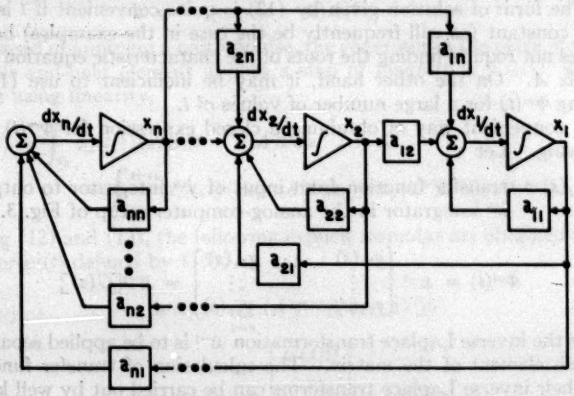


FIG. 3. An analog computer simulation.

In the absence of any external inputs to the continuous dynamic elements (13) we have by inspection of Fig. 3,

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{\gamma} a_{ij}x_j(t), \quad (i = 1, \cdots, \gamma). \tag{9-i}$$

This equation is a quantitative description of the state transition $x_i(t) \to x_i(t) + dx_i(t)$ due to the infinitesimal time change $t \to t + dt$. To calculate the state transition for a finite change in time, it is necessary to solve the system of differential equations (9). We first cast (9)

into the standard vector-matrix notation:

$$\frac{d\mathbf{x}^e(t)}{dt} = A\mathbf{x}^e(t) \tag{10}$$

where $\mathbf{x}^e(t)$ is a vector defined by (4) and A is a $(\gamma \times \gamma)$ constant matrix whose elements are defined above.

It is well known (14) that solutions of (10) starting at any time t_0 and observed at any time t are given by

$$\mathbf{x}^e(t) = (\exp\,(t - t_0)A)\mathbf{x}^e(t_0) = \Phi^{ee}(t - t_0)\mathbf{x}^e(t_0) \tag{11}$$

where

$$\exp tA = \sum_{k=0}^{\infty} t^k A^k/k! \quad (A^0 = \text{unit matrix}). \tag{12}$$

The infinite series (12) converges for all finite values of t. $\Phi^{ee}(t)$ is called the transition matrix (11) of the continuous dynamic element.

The form of solution given by (12) is quite convenient if t is some fixed constant (as will frequently be the case in the examples) because it does not require finding the roots of the characteristic equation of the matrix A. On the other hand, it may be inefficient to use (12) for finding $\Phi^{ee}(t)$ for a large number of values of t.

A convenient way of obtaining a closed expression for $\Phi^{ee}(t)$ is the following. Let

$g_{ij}(s)$ = transfer function from input of j^{th} integrator to output of i^{th} integrator in the analog-computer setup of Fig. 3.

Then

$$\Phi^{ee}(t) = \mathfrak{L}^{-1}\begin{bmatrix} g_{11}(s) & \cdots & g_{1\gamma}(s) \\ \vdots & & \vdots \\ g_{\gamma 1}(s) & \cdots & g_{\gamma\gamma}(s) \end{bmatrix} = \mathfrak{L}^{-1}[G(s)] \tag{13}$$

where the inverse Laplace transformation $\mathfrak{L}^{-1}$ is to be applied separately to each element of the matrix. The calculation of transfer functions and their inverse Laplace transforms can be carried out by well known methods of linear system engineering. In the examples which follow, calculation of $\Phi^{ee}(t)$ by means of (13) will therefore not require additional explanation.

The following important properties of $\Phi^{ee}(t)$ should be noted:

$$\Phi^{ee}(t_1 + t_2) = \Phi^{ee}(t_1)\Phi^{ee}(t_2) \quad \text{for all } t_1, t_2. \tag{14}$$

This follows directly from (12). Therefore, in particular, we have

$$\Phi^{ee}(0) = \text{unit matrix}; \quad (\Phi^{ee}(t))^{-1} = \Phi^{ee}(-t). \tag{15}$$

To complete the derivation of the transition equations for the continuous dynamic elements, consider now the inputs to these elements. As mentioned previously, the external inputs to these elements are

applied through sample-and-hold elements. (There may be inputs from one continuous dynamic element to another but these are implicitly taken care of in the definition of the A matrix in (9).)

When the sampling operations are nonsynchronous, multiple-order, multi-rate or random, the output of the sample-and-hold element is a constant between successive sampling instants. Therefore the inputs to the continuous dynamic elements will be a series of steps.

Let us replace, for a moment, the j^{th} sample-and-hold element by means of a generator whose output at time $t = t_k$ is a unit impulse. The effect of the impulse is to change instantaneously the state of the continuous dynamic elements from $\mathbf{x}^c(t_k)$ to $\mathbf{x}^c(t_k^+) = \mathbf{x}^c(t_k) + \Delta\mathbf{x}^c(t_k)$. Let us denote the change by the γ-vector ($\gamma \times 1$ matrix)

$$\mathbf{f}^j = \Delta\mathbf{x}^c(t_k).$$

In view of the linearity, the effect of the unit impulse on the succeeding time-variation of the state variables is expressed by

$$\mathbf{x}^c(t) = (\exp\,(t - t_k)A)\mathbf{f}^j, \quad t > t_k.$$

If instead of applying a unit impulse, the generator in the place of the j^{th} sample-and-hold element applies a step of magnitude $x_j(t_k^+)$, we get, again using linearity,

$$\Delta\mathbf{x}^c(t) = \int_{t_k}^{t} x_j(t_k^+)(\exp\,(t - u)A)\mathbf{f}^j du \quad t > t_k$$

$$= \int_0^{t-t_k} x_j(t_k^+)(\exp\,uA)\mathbf{f}^j du = x_j(t_k^+)\mathbf{h}^j(t), \quad t > t_k. \tag{16}$$

Using (12) and (13), the following explicit formulas are obtained for the vector $\mathbf{h}^j(t)$ defined by (16):

$$\mathbf{h}^j(t) = \sum_{k=0}^{\infty} (A^k t^{k+1}/(k+1)!)\mathbf{f}^j \tag{17}$$

$$= \mathfrak{L}^{-1}\left[\frac{G(s)\mathbf{f}^j}{s}\right]. \tag{18}$$

We can now write down the complete transition equations for continuous dynamic elements. Let t_k be the time at which any one of the sample-and-hold elements changes it state. Let t_{k+1} be the instant at which the next change of state of any one of the sample-and-hold element occurs. Then

$$x_i(t_k + \tau) = \sum_{j=1}^{\gamma} \phi_{ij}^{oc}(\tau)x_j(t_k) + \sum_{j=\gamma+\delta+1}^{n} h_i^j(\tau)x_j(t_k^+),$$

$$0 < \tau \leqq t_{k+1} - t_k \quad (i = 1, 2, \cdots, \gamma) \tag{19-i}$$

It remains to consider sample-and-hold elements in which the sampling operation is of the noninstantaneous type described on p. 406.

The output of such elements after $t_k + U$ (see (2-*B*)) is constant, so that the previous analysis applies. During the interval $t_k < t \leqq t_k + U$, the output of the sample-and-hold element is equal to the input. The input to the sample-and-hold element must consist of combinations of state variables of continuous dynamic elements and the system input $r(t)$; no state variables of discrete dynamic elements can appear.

Let the $x_n(t)$ be the state variable of the *only* sample-and-hold element of the noninstantaneous type described on p. 409. By (7), and bearing in mind the above remarks, we have

$$x_n(t) = s_{n1}x_1(t) + s_{n2}x_2(t) + \cdots + s_{n\gamma}x_\gamma(t) + s_{n,n+1}r(t) \quad (7\text{-}n)$$

during the time interval $t_k < t \leqq t_k + U$. We now form a matrix $\bar{A}$ by defining its elements as

$$\bar{a}_{ij} = a_{ij} + f_i^n s_{nj}.$$

Similarly, we define a vector $\mathbf{f} = s_{n,n+1}\mathbf{f}^n$. Then the transition equations in the interval $t_k < t \leqq t_{k+1} + U$ are obtained by solving the differential equation

$$d\mathbf{x}^c/dt = \bar{A}\mathbf{x}^c + \mathbf{f}r(t) \qquad (20)$$

which yields

$$\mathbf{x}^c(t_k + \tau) = (\exp t\bar{A})\mathbf{x}^c(t_k) + \int_0^\tau r(u)(\exp(\tau - u)\bar{A})\mathbf{f}du, \qquad 0 < \tau \leqq U. \quad (21)$$

An analogous procedure may be employed when there is more than one noninstantaneous sampling element. The tedious, but elementary, details of setting up equations in those cases are left to the interested reader.

SYSTEM TRANSITION EQUATIONS

The system transition equations can now be written down very simply. The state transitions of each one of the three types of elements are represented by a matrix acting on the combined state variables. By multiplying together the various transition matrices, we obtain the transition matrix for the system for any two *successive* values of time.

We assume first that $r(t) = 0$.

Sample-and-Hold Elements

Let S_i $(i = \gamma + \delta + 1, \cdots, n)$ be the matrix obtained by replacing the i^{th} row of the $n \times n$ unit matrix by $s_{i\alpha}, s_{i\beta}, \cdots, s_{i,\gamma+\delta}, 0, \cdots, 0$. (See (7-*i*).)

$$S_i = \left[\begin{array}{c|c|c} \begin{matrix}1\end{matrix} & \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}0\end{matrix} \\ \hline \begin{matrix}0\end{matrix} & \begin{matrix}1 & & 0\end{matrix} & \begin{matrix}0\end{matrix} \\ \hline \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix} \\ \hline \begin{matrix}s_{i1} & \cdots & s_{i\gamma}\end{matrix} & \begin{matrix}s_{i,\gamma+1} & \cdots & s_{i,\gamma+\delta}\end{matrix} & \begin{matrix}0 & \cdots & 0 & \cdots & 0\end{matrix} \\ \hline \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix} \end{array}\right] \begin{matrix}\\ \\ \\ \leftarrow i^{\text{th}} \text{ row} \\ \\ \\ \end{matrix} \qquad (22\text{-}i)$$

If the state $\mathbf{x}(t_k)$ of the over-all system is multiplied by the $n \times n$ matrix S_i, then all states remain the same except the i^{th} which changes in accordance with (7-i). Note also the relations

$$S_iS_j = S_jS_i \qquad \text{and} \qquad S_iS_i = S_i.$$

Discrete Dynamic Elements

Let $D_{ij\cdots}(i, j, \cdots = \gamma + 1, \cdots, \gamma + \delta)$ be the matrix obtained by replacing the i^{th} row of the $n \times n$ unit matrix by $d_{i1}, \cdots, d_{i,\gamma+\delta}, 0, \cdots, 0$; replacing the j^{th} row by $d_j, \cdots, d_{j,\gamma+\delta}, 0, \cdots, 0$; etc. (See (8-$i$), 8-$j$)):

$$D_{ij\cdots} = \left[\begin{array}{c|c|c} \begin{matrix}1\end{matrix} & \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}0\end{matrix} \\ \hline \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}1 & \cdots & 0\end{matrix} & \\ \begin{matrix}d_{i1} & \cdots & d_{i\gamma}\end{matrix} & \begin{matrix}d_{i,\gamma+1} & \cdots & d_{i,\gamma+\delta}\end{matrix} & \begin{matrix}0\end{matrix} \\ \begin{matrix}d_{j1} & \cdots & d_{j\gamma}\end{matrix} & \begin{matrix}d_{j,\gamma+1} & \cdots & d_{j,\gamma+\delta}\end{matrix} & \\ \begin{matrix}0 & \cdots & 0\end{matrix} & \begin{matrix}0 & \cdots & 1\end{matrix} & \\ \hline \begin{matrix}0\end{matrix} & \begin{matrix}0\end{matrix} & \begin{matrix}1\end{matrix} \end{array}\right] \begin{matrix}\\ \\ \leftarrow i^{\text{th}} \text{ row} \\ \leftarrow j^{\text{th}} \text{ row} \\ \\ \\ \end{matrix} \qquad (23\text{-}ij\cdots)$$

If the state $\mathbf{x}(t_k)$ of the over-all system is multiplied by the $n \times n$ matrix $D_{ij\cdots}$ then all the states remain the same, except $x_i(t)$, $x_j(t)$, $\cdots$ which undergo a transition in accordance with (8-i), (8-j), $\cdots$, all at the *same* time $t_i^+ = t_j^+ = \cdots$.

Continuous Dynamic Elements

Let $\Phi(\tau)$ be the matrix defined by:

$$\Phi(\tau) = \left[\begin{array}{c|c|c} \Phi^{ee}(\tau) & 0 & \begin{matrix}h_1^{\gamma+\delta+1}(\tau) & \cdots & h_1^{n}(\tau) \\ \vdots & & \vdots \\ h_\gamma^{\gamma+\delta+1}(\tau) & \cdots & h_\gamma^{n}(\tau)\end{matrix} \\ \hline \begin{matrix}0\end{matrix} & \begin{matrix}1\end{matrix} & \begin{matrix}0\end{matrix} \\ \hline \begin{matrix}0\end{matrix} & \begin{matrix}0\end{matrix} & \begin{matrix}1\end{matrix} \end{array}\right] \begin{matrix}\\ \\ \}\text{first } \gamma \text{ rows} \\ \\ \\ \end{matrix} \qquad (24)$$

If the state $\mathbf{x}(t)$ of the over-all system is multiplied by the $n \times n$ matrix $\Phi(\tau)$, then the states of the discrete dynamic elements and sample-and-hold elements remain the same and the states of the continuous dynamic elements undergo a continuous transition as a function of τ. The argument of $\Phi(\tau)$ may have any value in the interval $0 < \tau \leqq t_{k+1} - t_k$, where t_k is the last sampling instant and t_{k+1} is the next sampling instant.

The system transition matrix, for any interval of time, is formed by multiplying together the individual matrices S_i, D_{ij}... and $\Phi(\tau)$ as required by the various sampling operations taking place in the system.

Effect of System Input

We now remove the restriction $r(t) = 0$. For simplicity, we assume that $r(t)$ can affect the system only through a sample-and-hold element. This will exclude minor complications due to the convolution integral in (21). In accordance with the previous generalized notation, we now define the n-vectors

$$\mathbf{d}_i = \begin{bmatrix} 0 \\ \hline 0 \\ \vdots \\ d_{i,n+1} \\ \vdots \\ 0 \\ \hline 0 \end{bmatrix} \qquad \mathbf{s}_i = \begin{bmatrix} 0 \\ \hline 0 \\ \hline 0 \\ \vdots \\ s_{i,n+1} \\ \vdots \\ 0 \end{bmatrix} \tag{25}$$

which correspond to the coefficients of $r(t)$ on the right-hand sides of (7-i) and (8-i). These vectors show the effect of $r(t)$ on the state transitions.

Example 1. System Defined by Its Transition Equations

The over-all transition equations of a given system have the form:

$$\mathbf{x}(t_{k+1}) = \Phi(T_B)S_{10}\Phi(T_A)D_{789}S_{11}\mathbf{x}(t_k) + \Phi(T_B)S_{10}\Phi(T_A)(\mathbf{d}_7 + \mathbf{s}_{11})r(t_k). \tag{26}$$

The question is: What kind of a system is the one whose transition equation is given by (26)? The answer can be obtained by systematically examining the various terms in (26).

(a) There are eleven state variables. The first six state variables belong to continuous dynamic elements, the next three to discrete dynamic elements, the last two to sample-and-hold elements.

(*b*) Each sampling operation has the same period $T = T_A + T_B$; the sampling operations in connection with the discrete dynamic element and the second sample-and-hold element (whose output is x_{11}) occur simultaneously at t_k; the sampling operation in connection with the first sample-and-hold element (whose output is x_{10}) occurs at $t_k' = t_k + T_A$. Thus the sampling is of the nonsynchronous, fixed-period type.

(*c*) All information regarding the input $r(t)$ is obtained at t_k. The first sample-and-hold element samples only the state variables of the continuous and discrete dynamic elements. This may happen, for instance, in order to utilize the computations performed by the discrete dynamic element (which may be assumed to require less time than T_A for completion) before the next sampling point t_{k+1}.

It is easy to see that (26) may represent the operation of a system such as Fig. 1.

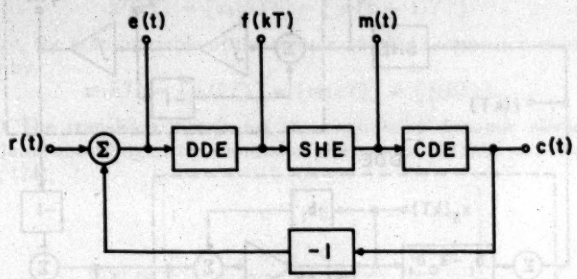


FIG. 4. Example 2.

SOME ILLUSTRATIVE EXAMPLES

To show how the formalism of the preceding sections is applied to the derivation of transition equations of systems containing various types of sampling operations, one example for each type of sampling will be given in this section.

In each example, the derivation of the transition equations requires essentially four steps:

Step (*i*). Selection of the state variables.

Step (*ii*). Calculation of the various matrices and vectors defined in the two previous sections.

Step (*iii*). Derivation of the transition equations for the various sampling operations and sample-free time-intervals.

Step (*iv*). Combination and simplification of the transition equations.

Example 2. Conventional Sampling System

A typical system of this class is shown in Fig. 4. The system con-

tains a discrete dynamic element (for improving the stability of the system). The sampling operations in the discrete dynamic element and the sample-and-hold element are performed synchronously. (This is not always necessary or even desirable in practice. See Kranc (6).) Thus sampling occurs at the instants:

$$\textit{Discrete Dynamic Element:}\quad 0, T, \cdots, kT, \cdots$$
$$\textit{Sample-and-Hold Element:}\quad 0, T, \cdots, kT, \cdots$$

(i) The continuous dynamic element in Fig. 4 is described by the differential equation

$$\frac{d^2c}{dt^2} + \frac{dc}{dt} = m(t) \tag{27}$$

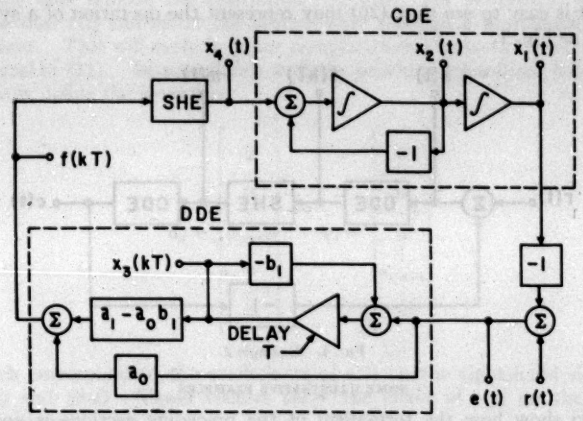


FIG. 5. Example 2 (details).

which relates the output $c(t)$ of the system to the control signal $m(t)$. The discrete dynamic element is described by the difference equation

$$f(kT) = a_0e(kT) + a_1e((k-1)T) - b_1f((k-1)T). \tag{28}$$

The input signal $e(t)$ to the discrete dynamic element is continuous. This signal is sampled and only its samples affect the state transitions in the discrete dynamic element. The output $f(kT)$ of the discrete dynamic element is a discrete signal, that is, it is available only at time kT.

The state variables for the continuous dynamic element are selected as follows:

$$\mathbf{x}^c(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c(t) \\ dc(t)/dt \end{bmatrix}. \tag{29}$$

With this choice of the state variables, the block diagram for the continuous dynamic element governed by (27) is shown in Fig. 5. This block diagram shows also how Eqs. 27 could be simulated by means of an analog computer.

To define the state variables for the discrete dynamic element, (28) is written in the form

$$f(kT) = a_0 e(kT) + (a_1 - a_0 b_1) y((k-1)T) \tag{30}$$

$$y(kT) = e(kT) - b_1 y((k-1)T). \tag{31}$$

By eliminating $y(kT)$ and $y((k-1)T)$ from (30) and (31), it follows easily that Eqs. 30 and 31 are equivalent to (28). From this, it is clear that the state of the discrete dynamic element at time kT is given by $y((k-1)T)$. Hence let

$$\mathbf{x}^d(kT) = [x_3(kT)] = [y((k-1)T)]. \tag{32}$$

Finally, the state variable of the sample-and-hold element is evidently given by:

$$\mathbf{x}^s(kT) = [x_4(kT)] = [m(kT)] = [f(kT)]. \tag{33}$$

(*ii*) The transition matrix for the continuous dynamic element is calculated according to the discussion given on p. 419; the end result is, using (24),

$$\Phi(\tau) = \left[\begin{array}{c|c|c|c} 1 & 1-e^{-\tau} & 0 & \tau - 1 + e^{-\tau} \\ \hline 0 & e^{-\tau} & 0 & 1 - e^{-\tau} \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]. \tag{34}$$

Using the fact that $e(t) = r(t) - x_1(t)$ and the notation of (32), we can write (30) and (31) in the form

$$x_3((k+1)T) = -x_1(kT) - b_1 x_3(kT) + r(kT)$$
$$f(kT) = (a_1 - a_0 b_1) x_3(kT) + a_0[r(kT) - x_1(kT)].$$

Hence, from (23) and (25)

$$D_3 = \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline -1 & 0 & -b_1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \tag{35}$$

$$\mathbf{d}_3 = \begin{pmatrix} 0 \\ 0 \\ \hline 1 \\ \hline 0 \end{pmatrix}. \tag{36}$$

If Eq. 30 is changed similarly, using the notation of (33), we get

$$x_4(kT^+) = -a_0x_1(kT) + (a_1 - a_0b_1)x_3(kT) + a_0r(kT).$$

Hence, from (22) and (25),

$$S_4 = \left(\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline -a_0 & 0 & a_1 - a_0b_1 & 0 \end{array}\right) \tag{37}$$

$$\mathbf{s}_4 = \begin{pmatrix} 0 \\ 0 \\ \hline 0 \\ \hline a_0 \end{pmatrix}. \tag{38}$$

(iii–iv) To obtain the transition equations for the over-all system, we must examine the sequence of the various transitions. Consider the time $t = kT$. First, the new value of the control signal $x_4(kT)$ is computed and we have

$$\mathbf{x}(kT^+) = S_4\mathbf{x}(kT) + \mathbf{s}_4r(kT).$$

Next, the data for the computation of the state transition of the discrete dynamic element are collected and the new state x_3 is calculated.

$$\begin{aligned} \mathbf{x}(kT^{++}) &= D_3\mathbf{x}(kT^+) + \mathbf{d}_3r(kT) \\ &= D_3S_4\mathbf{x}(kT) + (\mathbf{d}_3 + \mathbf{s}_4)r(kT). \end{aligned} \tag{39}$$

Actually, the new value of x_3 may not be available at kT^{++} but only

somewhat later; this is immaterial, however, as far as the transition equations are concerned.

During the interval $kT < t \leqq (k+1)T$ the only transitions taking place are those of the state variables of the continuous dynamic element. Hence,

$$\mathbf{x}(kT+\tau) = \Phi(\tau)[D_3 S_4 \mathbf{x}(kT) + (\mathbf{d}_3 + \mathbf{s}_4) r(kT)], \quad 0 < \tau \leqq T \tag{40}$$

$$= \Psi(\tau)\mathbf{x}(kT) + \mathbf{v}(\tau) r(kT), \quad 0 < \tau \leqq T \tag{41}$$

where

$$\Psi(\tau) = \Phi(\tau) D_3 S_4 =$$

$$= \left[\begin{array}{cc|c|c} 1-a_0(\tau-1+e^{-\tau}) & 1-e^{-\tau} & (\tau-1+e^{-\tau})(a_1-a_0 b_1) & 0 \\ -a_0(1-e^{-\tau}) & e^{-\tau} & (1-e^{-\tau})(a_1-a_0 b_1) & 0 \\ \hline -1 & 0 & -b_1 & 0 \\ \hline -a_0 & 0 & a_1-a_0 b_1 & 0 \end{array}\right] \tag{42-A}$$

and

$$\mathbf{v}(\tau) = \Phi(\tau)(\mathbf{d}_3 + \mathbf{s}_4) = \left[\begin{array}{c} a_0(\tau - 1 + e^{-\tau}) \\ a_0(1-e^{-\tau}) \\ \hline 1 \\ \hline a_0 \end{array}\right] \tag{42-B}$$

Equation 41 describes the behavior of the system at any instant of time.

Remark b. By letting $\tau = T$ in (41), we have a *stationary* (constant-coefficient) difference equation

$$\mathbf{x}((k+1)T) = \Psi(T)\mathbf{x}(kT) + \mathbf{v}(T) r(kT) \tag{43}$$

relating the state variables at succeeding sampling instants. This is equivalent to the result usually obtained by means of the z-transform method.

Remark c. The last column of the matrix $\Psi(\tau)$ contains only zeros. This shows that the value of x_4 need not be known at the k^{th} sampling instant in order to find the state vector at the $(k+1)^{\text{st}}$ sampling instant. Thus x_4 may be dropped in Eqs. 41 or 43. If x_4 had been dropped before obtaining the final result, however, one could not have expressed $\Psi(\tau)$ as a simple product of matrices.

Remark d. In deriving (41) and (43), no real use was made of the fact that $T_k = T = \text{const.}$ Hence these equations are valid also when the sampling period is not a constant. In that case the transitions between successive sampling instants are governed by the *nonstationary* (varying-coefficient) difference equation

$$\mathbf{x}(t_{k+1}) = \Psi(T_k)\mathbf{x}(t_k) + \nu(T_k)r(t_k) \tag{44}$$

where t_k denotes the k^{th} sampling instant and $T_k = t_{k+1} - t_k$.

Example 3. Multiple-Order Sampling

We consider the same system as in Example 2. We assume again that the sampling operations are synchronous, but now they take place at the instants

$$t_k = 0, \quad T_0, \quad T_0+T_1, \cdots, T_0+\cdots+T_{q-1}=T, \cdots, kT, \quad kT+T_0, \cdots.$$

In other words, the sampling period T_k is a periodic function of k with period q.

(i–iii) This example is evidently a special case of the nonstationary system (44) derived in connection with the previous example.

(iv) Because T_k is periodic, we may relate $x(kT)$ to $x((k+1)T)$ by means of the stationary difference equation:

$$\begin{aligned}\mathbf{x}((k+1)T) = {} & \Psi(T_0, \cdots, T_{q-1})\mathbf{x}(kT) \\ & + \nu_1(T_0, \cdots, T_{q-1})r(kT) \\ & + \nu_2(T_1, \cdots, T_{q-1})r(kT + T_0) + \cdots \\ & \qquad + \nu_q(T_{q-1})r(kT + T_0 + \cdots + T_{q-2})\end{aligned} \tag{45}$$

where

$$\begin{aligned}\Psi(T_0, \cdots, T_{q-1}) &= \Psi(T_{q-1})\Psi(T_{q-2}) \cdots \Psi(T_0) \\ \nu_1(T_0, \cdots, T_{q-1}) &= \Psi(T_{q-1})\Psi(T_{q-2}) \cdots \nu(T_0) \\ \nu_2(T_1, \cdots, T_{q-1}) &= \Psi(T_{q-1})\Psi(T_{q-2}) \cdots \nu(T_1) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \nu_q(T_{q-1}) &= \nu(T_{q-1}),\end{aligned} \tag{46}$$

where $\Psi(\tau)$, $\nu(\tau)$ are given by (42-*A*-*B*).

Example 4. Nonsynchronous Sampling

The system is shown in Fig. 6. The sampling operations occur at the following instants of time:

Sample-and-Hold Element 1: $0, \quad T, \cdots, \quad kT, \cdots$

Sample-and-Hold Element 2: $U, T + U, \cdots, kT + U, \cdots$

For simplicity, it is assumed that $T > U > 0$.

(i) The definition of the state vector is apparent from Fig. 6.

(ii) The transition matrices for the two sample-and-hold elements are given by

$$S_3 = \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \quad \text{and} \quad S_4 = \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 \end{array}\right]$$

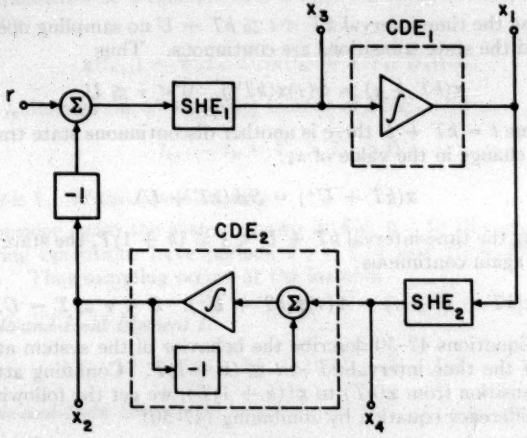


FIG. 6. Example 4.

while the transition matrix of the continuous element is

$$\Phi(\tau) = \left[\begin{array}{c|c|c|c} 1 & 0 & \tau & 0 \\ \hline 0 & e^{-\tau} & 0 & 1 - e^{-\tau} \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right].$$

From Fig. 6 we have also

$$\mathbf{s}_3 = \begin{bmatrix} 0 \\ 0 \\ \hline 1 \\ \hline 0 \end{bmatrix} \quad \text{and} \quad \mathbf{s}_4 = 0.$$

(iii) At time $t = kT$, there is a discontinuous change in the value of x_3 which is expressed by

$$\mathbf{x}(kT^+) = S_3\mathbf{x}(kT) + \mathbf{s}_3 r(kT). \tag{47}$$

During the time-interval $kT < t \leqq kT + U$ no sampling operations occur and the state transitions are continuous. Thus

$$\mathbf{x}(kT + \tau) = \Phi(\tau)\mathbf{x}(kT^+), \quad 0 < \tau \leqq U. \tag{48}$$

At time $t = kT + U$ there is another discontinuous state transition due to a change in the value of x_4:

$$\mathbf{x}(kT + U^+) = S_4\mathbf{x}(kT + U). \tag{49}$$

During the time-interval $kT + U < t \leqq (k + 1)T$, the state transitions are again continuous:

$$\mathbf{x}(kT + U + \tau) = \Phi(\tau)\mathbf{x}(kT + U^+), \quad 0 < \tau \leqq T - U. \tag{50}$$

(iv) Equations 47–50 describe the behavior of the system at every instant of the time interval $kT < t \leqq (k + 1)T$. Confining attention to the transition from $\mathbf{x}(kT)$ to $\mathbf{x}((k + 1)T)$, we get the following stationary difference equation by combining (47–50)

$$\mathbf{x}((k + 1)T) = \Psi(T, U)\mathbf{x}(kT) + \nu(T)r(kT) \tag{51}$$

where

$$\Psi(T, U) = \Phi(T - U)S_4\Phi(U)S_3 =$$

$$= \left[\begin{array}{cc|cc} 1 & -T & 0 & 0 \\ 1 - e^{-(T-U)} & e^{-T} - U(1 - e^{-(T-U)}) & 0 & e^{-(T-U)} - e^{-T} \\ \hline 0 & -1 & 0 & 0 \\ 1 & -U & 0 & 0 \end{array}\right] \tag{52}$$

and

$$\mathbf{v}(T, U) = \Phi(T - U)S_4\Phi(U)\mathbf{s}_3 = \begin{bmatrix} T \\ U(1 - e^{-(T-U)}) \\ \hline 1 \\ U \end{bmatrix}.$$

Remark e. Notice that in (51) the state variable x_3 (but not x_4) is superfluous.

Remark f. As in Example 3, the derivation of the above equation does not make use of the assumption that T, U are constants. Letting these quantities be a function of k we get the nonstationary difference equation

$$\mathbf{x}(t_{k+1}) = \Psi(T_k, U_k)\mathbf{x}(t_k) + \mathbf{v}(T_k, U_k)r(t_k) \tag{53}$$

where t_k denotes the k^{th} sampling instant of SHE_1, and

$$t_{k+1} - t_k = T_k > U_k > 0.$$

Example 5. Multi-Rate Sampling

Consider again the system shown in Fig. 6. In this example, the sampling operations have periods $T_A = T/3$ and $T_B = T/2$, respectively. Thus sampling occurs at the instants

Sample-and-Hold Element 1:

$$kT, \quad kT + T/3, \quad kT + 2T/3, \quad (k + 1)T, \cdots$$

Sample-and-Hold Element 2:

$$kT, \quad kT + T/2, \quad (k + 1)T, \cdots$$

(i–ii) The transition matrices of the two sampling operations and of the continuous dynamic element are the same as in Example 4.

(iii) At time $t = kT$, both sample-and-hold elements have a discontinuous state transition:

$$\mathbf{x}(kT^+) = S_3S_4\mathbf{x}(kT) + \mathbf{s}_3r(kT).$$

During the interval $kT < t \leqq kT + T/3$ the state transitions are continuous

$$\mathbf{x}(kT + \tau) = \Phi(\tau)\mathbf{x}(kT^+), \quad 0 < \tau \leqq T/3.$$

At time $k = kT + T/3$,

$$\mathbf{x}(kT + T/3^{+}) = S_3\mathbf{x}(kT + T/3) + \mathbf{s}_3 r(kT + T/3).$$

During the interval $kT + T/3 < t \leqq kT + T/3$,

$$\mathbf{x}(kT + T/3 + \tau) = \Phi(\tau)\mathbf{x}(kT + T/3^{+}), \quad 0 < \tau \leqq T/6.$$

At time $t = kT + T/2$,

$$\mathbf{x}(kT + T/2^{+}) = S_4\mathbf{x}(kT + T/2).$$

During the interval $kT + T/2 < \tau \leqq kT + 2T/3$,

$$\mathbf{x}(kT + T/2 + \tau) = \Phi(\tau)\mathbf{x}(kT + T/2^{+}), \quad 0 < \tau \leqq T/6.$$

At time $t = kT + 2T/3$,

$$\mathbf{x}(kT + 2T/3^{+}) = S_3\mathbf{x}(kT + 2T/3) + \mathbf{s}_3 r(kT + 2T/3).$$

During the interval $kT + 2T/3 < t \leqq (k + 1)T$,

$$\mathbf{x}(kT + 2T/3 + \tau) = \Phi(\tau)\mathbf{x}(kT + 2T/3^{+}), \quad 0 < \tau \leqq T/3.$$

(iv) Since the pattern of sampling operations repeats with a common period T, by combining the preceding equations we can obtain a stationary difference equation relating $\mathbf{x}((k + 1)T)$ to $\mathbf{x}(kT)$:

$$\mathbf{x}((k + 1)T) = \Psi\mathbf{x}(kT) + \nu_1 r(kT) + \nu_2 r(kT + T/3) + \nu_3 r(kT + 2T/3) \quad (54)$$

where

$$\begin{aligned} \Psi &= \Phi(T/3)S_3\Phi(T/6)S_4\Phi(T/6)S_3\Phi(T/3)S_3S_4 \\ \nu_1 &= \Phi(T/3)S_3\Phi(T/6)S_4\Phi(T/6)S_3\Phi(T/3)\mathbf{s}_3 \\ \nu_2 &= \Phi(T/3)S_3\Phi(T/6)S_4\Phi(T/6)\mathbf{s}_3 \\ \nu_3 &= \Phi(T/3)\mathbf{s}_3. \end{aligned} \quad (55)$$

Remark g. It should be carefully noted that the derivation of the stationary difference Eq. 54 was possible only because the ratio of sampling periods was rational. In fact, if $T_A/T_B = p/q$ (where p, q are relatively prime integers), then the sampling patterns have a least common period $qT_A = pT_B = T$ and therefore it is possible to relate the state $\mathbf{x}(kT)$ to $\mathbf{x}((k + 1)T)$ by means of a stationary difference equation. When T_A/T_B is irrational or, what is practically the same thing, the integers p, q are large, then a very large number of intervals have to be considered in order to write down the over-all stationary transition Eq. 54.

Remark h. Alternately, it is possible to write down a nonstationary difference equation of the type

$$\mathbf{x}((k+1)T) = \Psi(k)\mathbf{x}(kT) + \mathbf{v}(k; r(t)) \tag{56}$$

where T is *not* the least common period of T_A and T_B. In this case, the matrix $\Psi(k)$ and the vector $\mathbf{v}(k; r(t))$ must be calculated separately for each interval $kT < t \leqq (k+1)T$ by considering the various sampling instants and the various sampling-free subintervals as was done above. When T_A/T_B is irrational, that is, when the least common period of T_A and T_B is infinite, the matrix $\Psi(k)$ will have coefficients which vary in a quasi-random (nonperiodic) fashion with k. Therefore under such circumstances one may expect the sampling system to exhibit a behavior which would be intuitively interpreted as "random."

Example 6. Noninstantaneous Sampling

This system is shown in Fig. 7. The sampling operation is as defined by (2–*B*.)

(i) The definition of the state variables is apparent from Fig. 7.

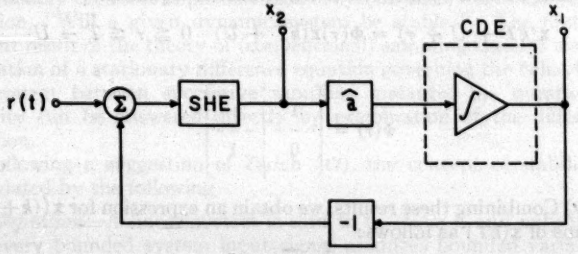


FIG. 7. Example 6.

(ii–iii) There are two transition matrices for the continuous dynamic element. During the time interval $kT < t \leqq kT + U$, the sample-and-hold element transmits all signals without modification. Thus

$$x_2(t) = r(t) - x_1(t). \tag{57}$$

During this time interval, the continuous dynamic element is governed by the differential equation

$$dx_1/dt = \bar{a}x_1 + r(t). \tag{58}$$

Integrating (58) and combining it with (57), we obtain the transition equation

$$\mathbf{x}(kT + \tau) = \bar{\Phi}(\tau)\mathbf{x}(kT) + \int_0^\tau \bar{\Phi}(\tau - \sigma)\bar{\mathbf{c}}_1 r(\sigma)d\sigma + \bar{\mathbf{s}}_2 r(kT + \tau), \qquad 0 < \tau \leqq U \quad (59)$$

where

$$\bar{\Phi}(\tau) = \left[\begin{array}{c|c} e^{\bar{a}\tau} & 0 \\ \hline -e^{\bar{a}\tau} & 0 \end{array}\right] \quad (60)$$

$$\bar{\mathbf{c}}_1 = \left[\begin{array}{c} 1 \\ \hline 0 \end{array}\right] \qquad \bar{\mathbf{s}}_2 = \left[\begin{array}{c} 0 \\ \hline 1 \end{array}\right].$$

During the time interval $kT + U \leqq t \leqq (k + 1)T$, the output of the sample-and-hold element remains constant and equal to $x_2(kT + U)$; the continuous dynamic element is governed by the differential equation

$$dx_1/dt = \bar{a}x_2(kT + U).$$

From this we find by inspection the transition equation

$$\mathbf{x}(kT + U + \tau) = \Phi(\tau)\mathbf{x}(kT + U) \quad 0 \leqq \tau \leqq T - U \quad (61)$$

where

$$\Phi(\tau) = \left[\begin{array}{c|c} 1 & \bar{a}\tau \\ \hline 0 & 1 \end{array}\right]. \quad (62)$$

(iv) Combining these results, we obtain an expression for $\mathbf{x}((k+1)T)$ in terms of $\mathbf{x}(kT)$ as follows:

$$\mathbf{x}((k + 1)T) = \Psi(T, U)\mathbf{x}(kT) + \nu(T, U, r(t)) \quad (63)$$

where

$$\Psi(T, U) = \Phi(T - U)\bar{\Phi}(U) = \left[\begin{array}{c|c} e^{\bar{a}U}(1 - \bar{a}(T - U)) & 0 \\ \hline -e^{\bar{a}U} & 0 \end{array}\right] \quad (64)$$

and

$$\nu(T, U, r(t)) = \Phi(T - U)\left[\mathbf{s}_2 r(kT + U) + \int_0^U \bar{\Phi}(U - \sigma)\bar{\mathbf{c}}_1 r(\sigma)d\sigma\right].$$

It should be observed that (63) is not a true difference equation (but rather a difference-integral equation) because of the complicated fashion in which the system input $r(t)$ affects the state transitions.

Example 7. Other Applications

There are many other problems in the analysis of nonstationary linear dynamic systems which can be handled by our methods. We mention two further examples:

(1) In discussing the solutions of Mathieu's equation with a square-wave parametric excitation, that is, the equation

$$d^2x/dt^2 + adx/dt + b(1 + c \operatorname{sgn}(\sin wt)) = 0$$

where a, b, c, w are positive constants and $\operatorname{sgn}(0) = 0$, $\operatorname{sgn} x = x/|x|$ if $x \neq 0$, Pipes (15) has used a method which is conceptually identical with ours.

(2) In a recent paper, Desoer (16) has discussed a dynamic behavior of networks which contain periodically operated switches. Desoer derived the equivalent of formula (12), apparently regarding this as a new result. In applying our method to Desoer's problem, very considerable conceptual and analytic simplifications may be obtained. We omit the details.

STABILITY

Probably the most important task of system theory is to answer the question, "Will a given dynamic system be stable?" The first significant result of the theory of (conventional) sampling systems was the derivation of a stationary difference equation governing the behavior of the system between successive sampling instants; the question of stability can be answered directly by examination of the difference equation.

Following a suggestion of Zadeh (17), the concept of stability is formulated by the following

> *Definition*—A linear system is said to be *stable* if and only if every bounded system input signal produces bounded variations in all state variables.

This definition applies to linear systems regardless of whether they are governed by differential or difference equations or whether they are stationary or nonstationary.

In a linear system, stability depends only on the transition matrix of the system; stability cannot be brought about or destroyed by a particular choice of the initial state or the system input signal. It is not difficult to show that the above definition implies.

> *Theorem 1. A stationary linear system is stable if and only if every element of* Ψ^N *tends to* 0 *as* $N \to \infty$ [*every element of* $\Psi(k)\Psi(k+1)\cdots\Psi(k+N)$ *tends to* 0 *uniformly*[4] *in* k *as* $N \to \infty$].

[4] It can easily be shown that if $\Psi(k)\Psi(k+1)\cdots\Psi(k+N)$ does not approach 0 *uniformly* as $N \to \infty$, then for some bounded system input signal some state variable will become unbounded.

In the above theorem, $\Psi(k)$ is the transition matrix of the system over some well-defined, constant interval of time.

One can "experimentally" verify from Theorem 1 the stability of a given system by multiplying together the transition matrix a large number of times. This is admittedly a brute-force procedure, but, because of the simplicity of matrix multiplication, very well suited for machine computation. Moreover, in the case of a nonstationary system, this method may be the only way of checking stability since no explicit necessary *and* sufficient stability conditions for nonstationary linear systems are known at present. (There are, of course, many sufficient and many necessary conditions.)

It is fortunate that in many sampling systems the transition matrices over sufficiently long intervals of time are constant; see Eqs. 43, 45, 51, 54 and 63. (For cases where the transition matrix is not constant, see Eqs. 44, 53, 56 and Remark g). For a constant transition matrix, the explicit stability conditions are as follows:

Theorem 2. A stationary linear system is stable if all n (possibly complex) zeros, z_i, *of the polynomial* $\det(\Psi - zI)$ *satisfy*

$$|z_i| < 1 \quad (i = 1, \cdots, n). \tag{65}$$

If, moreover, $|z_i| \neq 1$ *for any* i, *then condition (65) is not only sufficient but also necessary.*

The proof of this theorem follows readily by reducing Ψ to its Jordan canonic form.

Condition (65) is identical with the result that the poles z_i of the z-transform of the input-output relations of the system must lie within the unit circle (1). Of course, our result is much more general than the z-transform result because Ψ can be computed in many cases where the z-transform method is not applicable.

Finally, it should be remarked that a stationary transition matrix can be obtained in those cases and in those only where the pattern of the sampling operations repeats in a periodic fashion. This observation generalizes Floquet's theorem concerning the nonstationary differential equation

$$d\mathbf{x}/dt = A(t)\mathbf{x} \tag{66}$$

where $A(t) = A(t + T)$, that is, the system is periodic with period T. According to this theorem (14) the transition matrix $\Psi(t + T; t)$ of (66) for an interval of length T,

$$\mathbf{x}(t + T) = \Psi(t + T; t)\mathbf{x}(t)$$

is given by

$$\Psi(t + T; t) = \exp BT$$

where B is a constant matrix. This shows that the stability problem in a nonstationary but periodic linear system is essentially the same as in a stationary linear system provided that attention is focused on the transition matrix corresponding to one period.

CONCLUSIONS

The development of the theory and engineering application of sampling systems has been hampered by the fact that hitherto standard methods of analysis (the z-transform method and its various modifications) are not readily applicable unless all sampling operations in the system are synchronous and of constant sampling period. This paper shows that such restrictions are not at all basic; sampling operations of any presently conceivable kind can be readily studied within the unified framework presented here.

Sampling systems are a special class of nonstationary dynamic systems. As a result of the nonstationarity, whose complexity is directly dependent on the complexity of the sampling operations, the description of such systems is necessarily somewhat involved. In many cases the behavior of sampling systems is still stationary if attention is focused not on the instantaneous state transitions but on the transitions over a certain time-interval. The transition matrix for these intervals can then be used to answer the most basic question of system theory, namely that of stability.

In fact, any other information concerning the system behavior can be derived from the transition equations. Thus, the classical results concerning the stability of sampling systems can be extended to cases where the sampling operations are not restricted to the synchronous, constant-sampling-period case.

REFERENCES

(1) J. R. RAGAZZINI AND G. FRANKLIN, "Sampled-Data Control Systems," New York, McGraw-Hill Book Co., Inc., 1958.

(2) W. K. LINVILL, "Sampled-Data Control Systems Studied Through Comparison of Sampling with Impulse Modulation," *Trans. AIEE*, Vol. 70, pp. 1779–99 (1951).

(3) J. R. RAGAZZINI AND L. A. ZADEH, "The Analysis of Sampled-Data Systems," *Trans. AIEE*, Vol. 71, Part II, pp. 225–234 (1952).

(4) R. H. BARKER, "The Pulse Transfer Function and its Application to Sampling Servo Systems," *Proc. IEE*, Vol. 99, Part IV, pp. 302–317 (1952).

(5) J. G. TRUXAL, "Automatic Feedback Control System Synthesis," New York, McGraw-Hill Book Co., Inc., 1955, Chap. 9.

(6) G. KRANC, "Compensation of an Error-Sampled System by a Multi-Rate Controller," *Trans. AIEE*, Vol. 76, Part II, pp. 149–159 (1957).

(7) G. KRANC, "Additional Techniques for Sampled-Data Feedback Problems," IRE WESCON Convention Record, Vol. 1, Part IV, pp. 157–165, 1957.

(8) G. FARMANFARMA, "General Analysis and Stability Study of Finite Pulsed Feedback Systems," *Trans. AIEE*, Vol. 77, Part II, pp. 148–162 (1958).

(9) R. E. KALMAN AND J. E. BERTRAM, "General Synthesis Procedure for Computer Control of Single and Multi-Loop Linear Systems," *Trans. AIEE*, Vol. 77, Part II, pp. 602–609 (1958).

(10) R. E. KALMAN, "Optimal Nonlinear Control of Saturating Systems by Intermittent Action," IRE WESCON Convention Record, Vol. 1, Part IV, pp. 130–135 (1957).
(11) R. E. KALMAN AND R. W. KOEPCKE, "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indexes," *Trans. ASME*, Vol. 81, pp. 1820–1826 (1959).
(12) J. E. BERTRAM, "The Effect of Quantization in Sampled-Feedback Systems," *Trans. AIEE*, Vol. 77, Part II, pp. 177–182 (1958).
(13) R. E. KALMAN, "Analysis and Design Principles of Second and Higher Order Saturating Servomechanisms," *Trans. AIEE*, Vol. 74, Part II, pp. 294–310 (1955).
(14) E. A. CODDINGTON AND N. LEVINSON, "Theory of Ordinary Differential Equations," New York, McGraw-Hill Book Co., Inc., 1955, Chap. 3.
(15) L. A. PIPES, "Matrix Solution of Equations of the Mathieu-Hill Type," *J. Appl. Phys.*, Vol. 24, pp. 902–910 (1953).
(16) C. A. DESOER, "A Network Containing a Periodically Operated Switch Solved by Successive Approximations," *Bell System Tech. J.*, Vol. 36, pp. 1403–1428 (1957).
(17) L. A. ZADEH, "A General Theory of Linear Signal Transmission Systems," JOUR. FRANKLIN INST., Vol. 253, pp. 293–312 (1952).
(18) R. E. KALMAN, "Analysis and Synthesis of Linear Systems Operating on Randomly Sampled Data," doctoral thesis, Dept. of Electrical Engineering, Columbia University, New York, N. Y., 1957.
(19) R. E. KALMAN, "On a New Approach to Filtering and Prediction Problems," *ASME* paper 59–IRD–11; to appear in *J. Basic Engineering*, Vol. 1, 1959.
(20) FAN CHUN VUI, "On Pulsed Servo Systems Containing Two Pulse Elements with Unequal Pulse Repetition," (in Russian), *Automatika i Telemekhanika*, Vol. 19, pp. 917–930 (1958).
(21) YA. Z. TSYPKIN, "Theory of Pulse Systems," (in Russian), Moscow, State Press for Physico-Mathematical Literature, 1958.
(22) G. P. TARTAKOVSKIĬ, "On Pulse Systems with Variable Elements," *Elektrovias* 1956; *Radiotekhnika i Elektronika*, Vol. 2, pp. 15–22 (1957).

V.2
195

THE FRANKLIN INSTITUTE

MINUTES OF THE STATED MEETING

April 15, 1959

The Stated Meeting of The Franklin Institute was held at 8:15 P.M. in the Lecture Hall. Mr. Wynn Laurence LePage, President, called the meeting to order with approximately 225 members and guests in attendance.

The President stated the minutes of the March 18, 1959 meeting will be published in the April 1959 issue of the JOURNAL and will be presented for approval at the next meeting.

The President noted that this was the occasion of the Sixth Edward G. Budd lecture, which lectureship was founded in 1948 in honor and in memory of our distinguished former member of the Institute and its Board of Managers. Until his death in November 1946, Mr. Budd had been a member of the Institute since 1895 and a member of its Board of Managers since 1927.

The President then introduced the Budd Lecturer, Mr. Henry Belin duPont, Vice President of the E. I. duPont de Nemours & Company, who spoke on "The Greatest Invention of Them All."

Mr. duPont spoke of the real progress that has been made in technology, in the 135 years since The Franklin Institute was founded, noting Philadelphia's rich technology history. He recounted the difficulties faced by early inventors, not only in getting backing for their ideas but in actually creating inventions, for often they could only make a part-time job of it while they made their living at something else. He brought out that the modern industrial establishment has made the pursuit of technology a full-time job, not only for a few but for hundreds of thousands. Moreover, the development of the modern corporation has helped to increase and broaden the resources. It has provided a focal point for the development of technology and, while the research efforts of the universities, private foundations and government laboratories contribute importantly to the advancement of American science and engineering, these institutions could not exist were it not for the wealth generated by American industry.

Following Mr. duPont's talk there was an active and most interesting question period. Before adjourning the meeting at 9:20 P.M., the President expressed on behalf of the Institute its warm thanks to Mr. duPont.

WILLIAM F. JACKSON, JR.
Secretary

COMMITTEE ON SCIENCE AND THE ARTS

(Abstract of Proceedings of Stated Meeting held Wednesday, April 8, 1959.)

HALL OF THE COMMITTEE,
PHILADELPHIA, APRIL 8, 1959.

DR. WILLIAM G. SCHMIDT in the Chair.

The following report was presented for final action:

No. 3333: Work of John Hays Hammond, Jr., in Remote and Automatic Controls.

This report recommended the award of an Elliott Cresson Medal to John Hays Hammond, Jr., of Gloucester, Massachusetts, "In consideration of his many inventions and developments in the fields of radio and electronics; and, particularly, for his pioneering contributions and leadership in the invention, design and development of remote radio control for vehicles moving on or under the water."

D. S. FAHRNEY
Secretary to Committee

YOU CAN ADVANCE SCIENCE EDUCATION

Today, more than ever before in its 135-year history, there is vital need for The Franklin Institute to effectively promote education in science and technology. It is imperative that we meet this challenge by providing adequate educational opportunities in these fields. This requires vision, objective planning, and money. We have more than enough of the first two requisites, but far too little of the third.

Our programs are aimed at professional scientists and industry, as well as the lay public and young people seeking inspiration and guidance in choosing a career. The Institute's educational programs are impressive, for they begin with students in the early grades of our elementary schools and continue throughout an individual's professional or industrial life. With more funds at our disposal, the scope and vigor of these activities could be greatly increased.

The Franklin Institute is not richly endowed. It is a non-profit organization, depending for encouragement and support on an understanding public. Capable and conservative management assures wise administration of all funds.

Your gift or bequest, large or small, will be deeply appreciated and will be used effectively to broaden the Institute's educational usefulness. There is a warm satisfaction in giving financial support to an organization that has pioneered in, and is dedicated to, the advancement of science and technology through education.

When property is transferred, title should be in the name of The Franklin Institute of the State of Pennsylvania for the Promotion of the Mechanic Arts.

The Secretary of The Franklin Institute will gladly furnish you with additional information. Write to him at The Franklin Institute, Benjamin Franklin Parkway at Twentieth Street, Philadelphia 3, Pennsylvania.

V.2(
195

MEMBERSHIP

The following elections were approved on April 2, 1959 by the Membership Committee by authority given to it by the Board of Managers.

ANNUAL

M. I. Arbogast
John J. Boericke, Jr.
Warren H. Brand
William S. Buzzard
Reuben Cardonick
Milton F. Chauner
Alighiero Cortese
Frank Hickman Cutaiar
Paul D. Dalsimer
Manfred DeRewal
Warren J. Donaldson
Howard Elliott Eldridge
John W Forney, Jr
Michael F. Gaughan
Grellett Gibbons-Neff
George E. Glover
Emil P. Gohn
Donald Cameron Hankin
John C. Hendricks
Edwin W. Hewitt
James F. Hinton
Barnett John Junker
Robert A. Kauffmann
Frank T. Kay
Otto M. Kershock
Sara J. Ketcham
David G Kilpatrick
Leonard G Ladof
John LaTreyte Lang
Edwin W. Lauterbach
Michele Levene
Joseph B. Levin, Jr.
A. Barton Lewis
Basil C. Lewis
Louie L. Lewis
Dudley P. Lighty, Jr
Eugene P. Lounsberry
John H. Luff
James E. Lyons, Sr.
Robert C. Madden
R. S. Mandelkorn
George Masurat
Albert C. McCoy
Edward Charles Mesiti
Arthur Ostroff
Sidney H. Perlman
William O. Price
Helen Sharkey
George W. Schelling
Douglas H. Simpson
Ruth Sokoloff
Edwin A. Speakman
William N. Thomas
Kent C. Thornton
S. Visvanathan
Harry L. Voorhees
William P. Welch
Lawrence F. Wilcox
William V. Wilkes
Maurice L. Wolpert
Ralph W. Wright

EDUCATOR

Mrs. Jarvis H. Cooper
James D Hardy
Forrest L. Lenker
Frederic M. Livezey
Clayton D. Loughran

NON-RESIDENT

Samuel Briskin
James P. McGowan
Walter J. McMahon

NECROLOGY

Henry L. Adler '55
Earle N. Barber '55
James B. Brown, Jr. '54
Norman P. Hutson '55
Ludwig Kerst '51
F. J. Wilson '56
S. G. Windell '36

LIBRARY

The Committee on Library desires to add to the collections any technical works that members may wish to contribute. Contributions will be gratefully acknowledged and placed in the Library. Duplicates received will be transferred to other libraries as gifts of the donor. 617 items have been added the past month.

Photostat service. Photostat prints of any material in the collections can be supplied on request.

The Library and reading room are open from 9 A.M. until 5 P.M. on Mondays, Tuesdays, Thursdays, Fridays and Saturdays; and 2 P.M. until 10 P.M. on Wednesdays. Beginning June 1, Saturday hours will be from 9 A.M. until Noon.

RECENT ADDITIONS

AERONAUTICS

DWINNELL, JAMES H. Principles of Aerodynamics. 1949

GROVER, JOHN HERBERT HARDY. Aircraft Communications Systems. 1957.

MILES, JOHN W. The Potential Theory of Unsteady Supersonic Flow. 1959.

NORTH ATLANTIC TREATY ORGANIZATION. ADVISORY GROUP FOR AERONAUTICAL RESEARCH AND DEVELOPMENT. Material Properties Handbook. Vol. 1. 1958.

ARCHITECTURE AND BUILDING

ROUILLION, LOUIS AND RAMSEY, CHARLES GEORGE. Architectural Details. 1924.

ASTRONOMY

BOK, BARTHOLOMEUS JAN. The Astronomer's Universe. 1958.

YOUNG, CHARLES AUGUSTUS. The Sun. 1897.

V.20
195

BIBLIOGRAPHY

U. S. AIR FORCE. CAMBRIDGE RESEARCH CENTER. GEOPHYSICS RESEARCH DIRECTORATE. A Bibliography of the Electrically Exploded Wire Phenomenon. Ed. 2. 1958.

CHEMISTRY AND CHEMICAL TECHNOLOGY

Advances in Chemical Physics. Vol. 1. 1958.

BAMFORD, C. H. AND OTHERS. The Kinetics of Vinyl Polymerization by Radical Mechanisms. 1958.

BASKERVILLE, CHARLES AND CURTMAN, LOUIS J. A Course in Qualitative Chemical Analysis. 1916.

COULSON, E. A. AND HERINGTON, E. F. G. Laboratory Distillation Practice. 1958.

FIESER, LOUIS FREDERICK AND FIESER, MARY A. Basic Organic Chemistry. 1959.

HOLLEY, CLIFFORD DYER. Analysis of Paint Vehicles, Japans and Varnishes. 1920.

INTERNATIONAL CONFERENCE ON BIOCHEMICAL PROBLEMS OF LIPIDS. Essential Fatty Acids. 1958.

MARTIN, F. S. AND MILES, G. L. Chemical Processing of Nuclear Fuels. 1958.

SCHÖNBERG, ALEXANDER. Präparative Organische Photochemie. 1958.

SMITH, JOSEPH MAUK AND VAN NESS, HENDRICK C. Introduction to Chemical Engineering Thermodynamics. 1959.

TÖDT, FRITZ. Elektrochemische Sauerstoffmessungen. 1958.

CIVIL ENGINEERING

URQUHART, LEONARD CHURCH. Civil Engineering Handbook. Ed. 4. 1959.

CRYSTALLOGRAPHY

KRISHNAN, R. S. Progress in Crystal Physics. Vol. 1. 1958.

ELECTRICITY AND ELECTRICAL ENGINEERING

STANEK, JOSEF. Technik Elektrischer Messgeräte. 1957.

ENGINEERING

BEEDLE, LYNN S. Plastic Design of Steel Frames. 1958.
HODGE, PHILIP GIBSON. Plastic Analysis of Structures. 1959.
IWIŃSKI, T. Theory of Beams. 1958.
LA LONDE, WILLIAM S. Professional Engineer's Examination Questions and Answers. 1956.
SINES, GEORGE AND WAISMAN, J. L., ED. Metal Fatigue. 1959.

FOOD

FRAZIER, WILLIAM CARROLL. Food Microbiology. 1958.

GEOLOGY

ATOMIC ENERGY (U.S.S.R.) SUPPL. 6. 1957. The Geology of Uranium. 1958.
DARWIN, CHARLES ROBERT. The Structure and Distribution of Coral Reefs. Ed. 3. 1897.

INDUSTRIAL DESIGN

BEL GEDDES, NORMAN. Horizons. 1932.
JOHNSON, PHILIP. Machine Art. 1934.
PEVSNER, NIKOLAUS. An Enquiry into Industrial Art in England. 1937.
READ, HERBERT. Art and Industry. 1934.

LUBRICATION

AMERICAN SOCIETY OF LUBRICATION ENGINEERS. Transactions. Vol. 1–2. 1958.

MANUFACTURE

HERCULES POWDER CO. CELLULOSE PRODUCTS DEPT. Nitrocellulose; Properties and Uses. 1955.
HOWARD, E. D., ED. Modern Foundry Practice. Ed. 3. 1959.

MATHEMATICS

ASELTINE, JOHN A. Transform Method in Linear System Analysis. 1958.
BERS, LIPMAN, ED. Contributions to the Theory of Partial Differential Equations. 1954.
DUNFORD, NELSON AND SCHWARTZ, JACOB T. Linear Operators. 1951.
EPSTEIN, L. IVAN. Nomography. 1958.
HEADY, EARL OREL AND CANDLER, WILFRED. Linear Programming Methods. 1958.
MORRIS, MAX AND BROWN, ORLEY E. Differential Equations. 1941.
SAATY, THOMAS L. Mathematical Methods of Operations Research. 1959.
WOODS, FREDERICK SHENSTONE AND BAILEY, FREDERICK H. A Course in Mathematics for Students of Engineering and Applied Science. Vol. 1–2. 1907.

MECHANICAL ENGINEERING

BUCKINGHAM, EARLE. Gear Ratio Tables for 4- 6- and 8-Gear Combinations. Ed. 1. 1958.
HINE, CHARLES R. Machine Tools for Engineers. Ed. 2. 1959.
KOVÁTS, ANDREAS DE AND DESMUR, GEORGES. Pumps, Fans and Compressors. 1958.
MARKS, LIONEL SIMEON, ED. Mechanical Engineers' Handbook. Ed. 6. 1958.

METALLURGY

BUNSHAH, ROINTAN FRAMROZE, ED. Vacuum Metallurgy. 1958.
INSTITUTION OF METALLURGISTS. Behavior of Metals at Elevated Temperatures. 1957.

METEOROLOGY

KOEPPE, CLARENCE EUGENE AND DELONG, GEORGE C. Weather and Climate. 1958.

MINERALOGY

DANA, EDWARD SALISBURY. A Text-Book of Mineralogy with an Extended Treatise on Crystallography and Physical Mineralogy. New Ed. 1905.

MISSILES AND ROCKETS

PUCKETT, ALLEN E. AND RAMO, SIMON. Guided Missile Engineering. 1959.
WARREN, FRANCIS A. Rocket Propellants. 1958.

NUCLEAR ENGINEERING

ETHERLINGTON, HAROLD, ED. Nuclear Engineering Handbook. 1958.
FINKELNBURG, KARL ERNST WOLFGANG. Einführung in die Atomphysik. Ed. 5 and 6. 1958.
HERTZ, GUSTAV, ED. Lehrbuch der Kernphysik. Vol. 1. 1958.
KOPFERMANN, HANS. Nuclear Moments. 1958.
MIALKI, WERNER. Kernverfahrenstechnik. 1958.

PHOTOGRAPHY

Jahrbuch der Fotografie. 1954.

PHYSICS

CROWHURST, NORMAN H. Audio Measurements. 1958.
CURTISS, LEON FRANCIS. An Introduction to Neutron Physics. 1959.
ECKERT, ERNST R. G. Heat and Mass Transfer. 1959.
ENDT, P. M. AND DEMEUR, M., ED. Nuclear Reactions. Vol. 1. 1959.
GALBRAITH, WILLIAM. Extensive Air Showers. 1958.
HIX, C. F. AND ALLEY, R. P. Physical Laws and Effects. 1958.
LEHNERT, B., ED. Electromagnetic Phenomena in Cosmical Physics. 1958.
LEIGHTON, ROBERT B. Principles of Modern Physics. 1959.
LYKOW, A. W. Transporterscheinungen in Kapillarporösen Körpern. 1958.
OFFICER, CHARLES B. Introduction to the Theory of Sound Transmission, with Application to the Ocean. 1958.
ROUSE, HUNTER, ED. Advanced Mechanics of Fluids. 1959.
SHPOLSKII, E. V. Atomphysik. Vol. 1–2. Ed. 2. 1957–58.
U. S. OFFICE OF NAVAL RESEARCH AND NATIONAL RESEARCH COUNCIL. Symposium on Naval Hydrodynamics. 1957.

PLASTICS

BERNHARDT, ERNEST G., ED. Processing of Thermoplastic Materials. 1959.
DUFFIN, DANIEL JOHN. Laminated Plastics. 1958.
SKEIST, IRVING. Epoxy Resins. 1958.
THOMPSON, MORTIMER STAFFORD. Gum Plastics. 1958.

RAILROADS

MENCKEN, AUGUST. The Railroad Passanger Car. 1957.

SCIENCE

GIDLEY, WILLIAM F. AND MORENO, JOE R., COMP. Prefixes and Suffixes of Scientific Nomenclature. Ed. 4. 1958.

STATISTICS

BLOM, GUNNAR. Statistical Estimates and Transformed Beta-Variables. 1958.

V.2(
195

THE FRANKLIN INSTITUTE LABORATORIES
NICOL H. SMITH, DIRECTOR

THERMOELECTRIC ENERGY CONVERSION IN SOLIDS

BY

NICHOLAS FUSCHILLO[1]

I. INTRODUCTION

We are presently in the midst of a period of great change in electrical technology. This change has hitherto been greatest in the field of electronics where a host of solid state phenomena and devices are finding application as circuit elements undertaking diverse tasks such as amplification, rectification, oscillation, and detection and interconversion of mechanical, electromagnetic and electrical impulses. To this formidable list has recently been added semiconductor thermoelectricity with its capability for the direct conversion of heat into electricity and the direct electrical transport of thermal energy from one place to another. This development could profoundly affect modern technology which is of course largely based on the interconversions of heat and electrical power. In semiconductor thermoelectric devices, these interconversions take place without recourse to mechanical machinery, and considerable simplifications as well as other obvious advantages can be achieved by the elimination of dynamos, electric motors, compressors, and working fluids in electrical power generation, refrigeration, and heating. The physical effects upon which these conversions of electrical and thermal energy are based (the effects of Peltier, Seebeck, Thomson, and Joule) have been well known for more than a century. Nevertheless, it was not until the rapid development of semiconductor physics in the last two decades that the conditions for maximization of the effects concerned were properly understood. This understanding was achieved through basic research in semiconductor physics and has led in the past few years to the emergence of these phenomena from the physicist's laboratory into the realm of highly promising engineering developments.

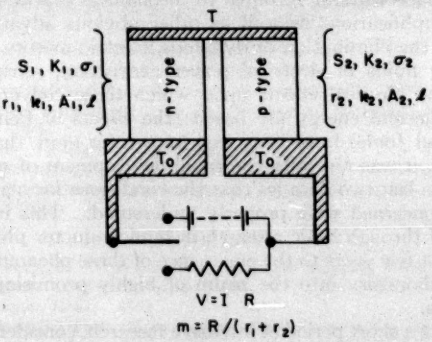
After quite a short period of intensive research, considerable progress has been made in the development of suitable materials for thermoelectric engines. Already existing materials are capable of power generating efficiencies approaching 10 per cent over useful temperature ranges, and efficiencies of 30 per cent may be achievable in a few years. In thermoelectric refrigeration materials, progress has also been encouraging. Figures of merit of 3.0×10^{-3} deg^{-1} at 300° K and temperature differences of 75° from 300°K have been obtained with single thermoelements. Moreover, the efficiency of thermoelectric heating can be made several times more efficient than ordinary electrical resistance heating due to the thermoelectric heat pumping action. Natur-

[1] Head, Magnetics and Semiconductors Branch.

ally, important problems still await solution. The figures of merit and efficiencies of heat pumping materials are too low by a factor of two or three for wide-spread conversion to thermoelectric methods in refrigeration and air conditioning. For thermoelectric generation, figures of merit of 3.0×10^{-3} deg^{-1} cannot as yet be maintained over a sufficiently large temperature to enable the attainment of energy conversion efficiencies competitive with the heat engine and dynamo combination. A number of additional problems also exist. Prominent among these is the contact resistance and bonding difficulties, which limit the miniaturization of thermoelectric devices, and the economical arrangement of effective heat exchange with the surroundings. We may expect, however, considerable progress to be made in the next few years in solving these problems in spite of their difficulty.

2. ENGINEERING RESULTS OF THE ZERO-ORDER PHENOMENOLOGICAL THEORY

The basic thermoelement (also known as a thermocouple) used as a thermoelectric engine is shown in the accompanying figure. Here, n-type and p-type semiconductor legs (unshaded) are connected by metallic current leads (shaded) so that a current, I, flows through them



in series. Referring to either leg 1 or 2 of the thermocouple we introduce the following notation; where the first four quantities represent average values over the appropriate temperature range.

S = Seebeck coefficient, volts/° K.
σ = specific electrical conductivity, ohm^{-1} cm^{-1}
κ = specific thermal conductivity, watts-cm^{-1} deg^{-1}
r = total electrical resistance, $r = r_1 + r_2$
k = total thermal conductance, $k = k_1 + k_2$
A = uniform cross section area of a leg
l = length of a thermocouple leg, $l_1 = l_2 = l$.

As n-type and p-type semiconducting materials are connected in series the Seebeck coefficient of the combination $S = S_1 - S_2$ will be additive so that its magnitude $|S| = |S_1| + |S_2|$. T and T_0 are the absolute temperatures of the variable and controlled (thermal sink) junctions of the semiconductor legs of the thermoelement. This temperature difference is caused by the reversible absorption and evolution of Peltier heat at T and T_0.

When used as a heat pump we may define the thermoelement efficiency by a coefficient of performance (C.O.P.) given by (see Ref. 1)

$$\mathcal{E} = \text{C.O.P.} = \frac{\text{Rate of absorption or evolution of heat at } T}{\text{Total electrical power consumed}} = \frac{\dot{Q}_T}{W}.$$

When the thermoelement is used as an electrical power generator, we consider the external current source replaced by an external circuit of load, R (shown dashed in the figure). The current, I, developed from the thermoelectric Seebeck e.m.f., is fed into the external circuit of impedence, R, where it does useful work. The useful power delivered to R will, of course, depend on the impedence match ratio $m = R/(r_1 + r_2)$.

The technical efficiency of the generator is defined as the ratio,

$$\eta = \frac{\text{Useful electrical power, } I^2R, \text{ delivered to the external circuit}}{\text{Rate of heat energy absorption at the variable junction at } T^\circ \text{ K.}}.$$

The principal results for thermocouples used as heat pumps and power generators with geometrically optimized areas are, as summarized in cases 1, 2 and 3 below provided we have:

$$\frac{A_1}{A_2} = \sqrt{\frac{\kappa_2\sigma_2}{\kappa_1\sigma_1}}; \quad N = \sqrt{1 + \tfrac{1}{2}Z_0(T + T_0)},$$

and the figure of merit,

$$Z = Z_0 = \frac{(S_1 - S_2)^2}{\left(\sqrt{\kappa_1/\sigma_1} + \sqrt{\kappa_2/\sigma_2}\right)^2}.$$

Case 1: Thermoelectric Refrigeration ($T < T_0$)

The optimum C.O.P. is given by the Carnot efficiency multiplied by a merit factor, M_C which must be less than unity due to irreversible dissipation processes such as thermal conduction,

$$\mathcal{E} = \mathcal{E}_{max} = \mathcal{E}_0 = \frac{T}{T_0 - T} \times M_C; \quad M_C = \frac{N - (T_0/T)}{N + 1}.$$

In addition $(T_0 - T)_{max} = \frac{1}{2}Z_0T$.

Case 2: Thermoelectric Heating $(T > T_0)$

$$\mathcal{E} = \mathcal{E}_{max} = \mathcal{E}_0 = \frac{T}{T - T_0} \times M_H; \quad M_H = 1 - \frac{2(N-1)}{Z_0T}.$$

Case 3: Thermoelectric Power Generation $(T >$ or $< T_0)$

The maximum technical efficiency, η_0, occurs for a value of $m = N$ and is given by the Carnot efficiency times a merit factor, M_G, which must be less than unity. Thus,

$$\eta_0 = \frac{T - T_0}{T} \times M_G; \quad M_G = \frac{N-1}{N + T_0/T}.$$

Among the most important features of these results are the following:

The coefficient of performance (efficiency) of a thermoelectric heat pump increases with decreasing temperature difference between the junctions, while the efficiency of the generator increases with increasing temperature difference. In both the heat pump and generator, the efficiency depends on a merit factor, M, which expresses the fraction of the thermodynamic or Carnot efficiency actually achieved. The extent to which M is less than unity is a measure of the efficiency loss due to irreversible energy exchanges such as Joule heat (I^2R) generation and thermal conduction.

In both generators and heat pumps, if the cross section areas of the semi-conductor legs have been adjusted to give optimum performance, then their η_0 or $\mathcal{E}_0$ values depend on a single parameter expressing the properties of the materials. This factor is called the figure of merit Z, and is written $Z_0 = Z = S^2\sigma/\kappa$, for a specific material, or a single leg of a thermocouple. In materials research for thermoelectric applications, it is thus Z which has to be maximized over the whole temperature range of application.

The C.O.P. $(\mathcal{E}_0)$ and generator (η_0) efficiencies are independent of the volume of semiconductor in the thermoelement. Thus the thickness l of the semiconductor legs may be decreased without altering $\mathcal{E}_0$ or η_0 until the semiconductor-metal contact resistance becomes a significant source of energy loss from Joule heating.

Some typical results of this phenomenological treatment for the temperature range covered by intermetallic semiconductor compounds are as shown in Tables I and II.

3. SOME TECHNICAL ADVANTAGES OF THERMOELECTRIC HEAT PUMPS AND GENERATORS

The prospects for widespread use of thermoelectric heat pumps are based essentially on certain characteristics that are summarized below:

TABLE I.—*Thermoelectric Heat Pump Performance at $T_0 = 300°$ K.*

Coefficient of Performance of a single stage thermoelectric heat pump used as a refrigerator or as a heater. $T_0 = 300°$ K.
ΔT (cooling) $= T_0 - T$, ΔT (heating) $= T - T_0$.

Temperature C°		5	10	20	25	30	40	50
difference F°		9	18	36	45	54	72	90
Figure of merit, $Z_0 = 5.0 \times 10^{-3}$ deg^{-1}								
C.O.P. ($\mathcal{E}_0$)	Cooling	12.4	6.1	2.7	2.0	1.6	1.0	0.7
	Heating	13.6	7.3	4.0	3.5	2.8	2.3	2.0
$Z_0 = 3.0 \times 10^{-3}$ deg^{-1}								
C.O.P. ($\mathcal{E}_0$)	Cooling	8.9	4.1	1.8	1.3	0.96	0.58	0.34
	Heating	9.7	5.2	3.0	2.5	2.2	1.8	1.6
$Z_0 = 2.5 \times 10^{-3}$ deg^{-1}								
C.O.P. ($\mathcal{E}_0$)	Cooling	7.6	3.3	1.5	1.0	0.78	0.44	0.22
	Heating	8.3	4.6	2.6	2.3	1.8	1.6	1.5
$Z_0 = 2.0 \times 10^{-3}$ deg^{-1}								
C.O.P. ($\mathcal{E}_0$)	Cooling	6.3	2.8	1.1	0.78	0.56	0.29	0.11
	Heating	7.0	4.0	2.3	2.0	1.7	1.5	1.25
$Z_0 = 1.5 \times 10^{-3}$ deg^{-1}								
C.O.P. ($\mathcal{E}_0$)	Cooling	4.9	2.2	0.83	0.55	0.33	0.12	...
	Heating	5.8	3.5	2.1	1.8	1.4	1.2	1.1

TABLE II.—*Thermoelectric Power Generation at Moderate Temperatures.*

Efficiency, η_0, of a single-stage generator. $T_0 = 300°$ K.

$$\Delta T = T - T_0, \; M = \eta_0 \frac{T}{\Delta T}.$$

Temperature C°	100	200	300	400	500	700
difference F°	180	360	540	720	900	1080
Figure of merit, $Z_0 = 5.0 \times 10^{-3}$ deg^{-1}						
η_0 %	6.8	12.5	17.2	21.5	25.0	32.0
Figure of merit, $Z_0 = 3.0 \times 10^{-3}$ deg^{-1}						
η_0 %	4.9	9.2	13.0	16.5	19.5	25.0
Figure of merit, $Z_0 = 2.0 \times 10^{-3}$ deg^{-1}						
η_0 %	3.8	7.1	10.0	13.0	15.5	20.0
Figure of merit, $Z_0 = 1.5 \times 10^{-3}$ deg^{-1}						
η_0 %	3.0	5.8	8.2	10.5	12.7	16.5

Thermoelements possess extreme technical simplicity. The absence of motors and compressors results in a minimum of maintenance and the absence of lubrication problems, fluid working substance, and noise or vibration. They thus possess high technical reliability combined with ease of automatic control and continuous variation of heat pumping capacity by simple current regulation.

Only a simple current reversal is required to reverse the direction of heat pumping to provide heating in cooling installations and vice versa. Thus convenient excursions can be made from room temperature in either direction with a single miniaturized thermoelement or a large battery of thermoelements.

Thermoelements have high coefficients of performance for applications requiring small temperature differences (Table I).

Thermoelectric heat pumps have greater efficiency than compressor units for small installations. Their efficiency will become comparable with compressor units for larger installations when Z values above 5×10^{-3} deg^{-1} are obtained.

Considerable adaptability in design is possible with heat pumping thermoelements; thus, panels of thermobatteries or individual thermoelements may be easily integrated with the most efficient heat exchange regions.

The first order theory shows that the C.O.P. of the thermoelectric heat pump is independent of the volume of semiconductor material. Thus, only a small thickness of semiconductor may eventually be required in thermoelements providing contact resistances can be sufficiently reduced.

Thermoelectric heat pumps can work effectively under extremes of ambient temperature or in corrosive environments where mechanical units would not function easily or reliably, for example, arctic, tropical, outer space, or missile environments.

Thermoelement fabrication should lend itself to inexpensive mass production techniques, since polycrystalline materials can be used with doping additions controlled to only moderate precision.

The chemical elements involved in many suitable materials are often inexpensive and are readily obtainable in large quantities.

An inherent disadvantage of thermoelectric heat pumps in household applications is the need for d.c. in their operation. An inexpensive germanium rectifying unit with low ripple content is thus necessary. On the other hand, the d.c. requirement is not a disadvantage in transport vehicles, or in laboratory and often factory installation, where d.c. is normally available.

The advantageous features of thermoelectric generation are similar to those summarized for the thermoelectric heat pump. In the thermoelectric generator, however, the efficiency increases with increasing temperature difference. This characteristic often offsets the greater diffi-

culty in obtaining high figures of merit with the presently investigated materials at temperatures in excess of 500° K. The reader is referred to Table II for the relative importance of these two factors. The most favorable characteristics of the thermoelectric generator follow, of course, from the elimination of the mechanical heat engine, and the direct and adaptable way in which thermoelements can be integrated into and around the heat source. Direct energy conversion is thus possible at the high temperature of the heat source, in contrast to the steam turbine and dynamo combination. Thermoelectric generation panels may be applied directly to large areas of a heated surface or thermoelements can be suitably integrated into pipes or baffles for heat exchange.

The thermoelectric material used will of course depend on the temperature range to be utilized. Naturally, a thermoelectric generator may advantageously consist of two or even three different materials suitable in different temperature ranges. Different series connected materials, each maximized by proper doping additions, could form composite thermocouple legs. The different portions of each leg would then generate electricity at optimized efficiencies in different temperature ranges, such as 2000 to 1200, 1200 to 700, and 700 to 300° K. As the temperature of the source, or combustion gas, decreased along the length of the heat exchange path, the higher temperature materials would be omitted in the thermocouples. Thus, such thermoelements would consist of successively three, two, and single-component semiconductor legs. Materials used for the range below 700° K would be generally similar to those employed in thermoelectric heat pumps operating near 300° K. A figure of merit of 3.0×10^{-3} deg^{-1} gives a calculated generator efficiency (η_0) of 17 per cent with a temperature difference of 400° above $T_0 = 300°$ K.

When discussing thermoelectric generation, the point is sometimes raised that each thermoelement develops only a very small voltage. However, it should be kept in mind, when comparing the thermoelectric generator and dynamo, that a single thermoelement, or a small group of thermoelements, is analogous in function to a single inductive turn on an electromagnetic generator. Thus a thermoelement, having a Seebeck coefficient $S = 0.24$ mV/° K., with $T = 800°$ K., and $T_0 = 300°$ K., generates 0.12 volt. If the thermoelement occupies an area of 2.0 cm^2, then a thermoelectric panel of area 1.0 m^2 should generate 600 volts. With good heat exchange arrangements, a volume of only 1.0 m^3 containing thermoelement arrays may yield ten to twenty times this voltage in a series connected arrangement. When such an array of thermojunctions is connected to produce maximum power in an external working load, a generator efficiency (η_0) of 18 per cent is obtained for a figure of merit of 2.5×10^{-3} deg^{-1}, with a temperature difference of 500° K. from 300° K.

Admittedly the thermoelectric technical efficiency possible with the materials measured to date is not equal to that achievable by conventional means, except in quite small capacity applications. It should be stressed, however, that technical efficiency is, after all, only one of the factors which decide the more relevant end product efficiencies such as saleability, cost, weight-space, and convenience. Theremoelectric methods already present many novel and attractive features as well as new design possibilities which are not feasible with conventional techniques of heat pumping and power generation. This factor can frequently offset the former disadvantage. Naturally these characteristic advantages will make thermoelectric machines even more attractive for larger capacity applications if their efficiency equals or exceeds electromechanical engines.

4. FIELDS OF APPLICATION

The characteristic advantages of thermoelectric engines lead to the promise of applications in many fields of industrial technology. We shall summarize some of these fields in broad categories in each of which exists a large number of different applications.

Heat Pumping

(*a*) Domestic, catering, and industrial refrigeration.
(*b*) Air conditioning and heating.
(*c*) Water cooling, drinking fountains.
(*d*) Numerous novel small household appliances involving small scale heating and cooling, for beverages, food and laundering.
(*e*) Household and hot water heating.
(*f*) Special cooling and heating applications to aeronautical, marine, missile and power reactor technology.
(*g*) Numerous chemical, agricultural, medical, dental, and scientific instrument applications.
(*h*) Thermostating and temperature control of electric components and circuits.
(*i*) Air conditioning, cooling, and heating applications in transportation.
(*j*) New fields of heat pump application arising out of the unique features of thermoelements.

Power Generation

(*a*) Nuclear power plant generation.
(*b*) Utilities and waste heat utilization.
(*c*) Generation of electricity from natural gas.
(*d*) Transportation: thermoelectrically-powered electric motor units

for propelling ships, submarines, locomotives, etc., using nuclear or hydrocarbon generated heat.

(*e*) Exhaust or waste heat power generation in transport vehicles (automobiles, tractors, locomotives, ships, submarines, aircraft, missiles), for battery charging, ignition, motor driving thermoelectric cooling, instrument and radio powering, etc.

(*f*) Solar generation.

(*g*) Portable generators of moderate power (less than 2000 watts) for isolated situations.

(*h*) Portable, very low power generators (less than 100 watts) for communication and other electronic applications.

5. PRESENT SITUATION REGARDING FIGURES OF MERIT

The results shown in Tables I and II stress the importance of fabricating materials of high figure of merit. In the low temperature range, −50 to 300° C., the best materials so far reported lie in the V–VI and IV–VI systems of semiconductor compounds. Solid solutions of the compounds within a system are often used to reduce the thermal conductivity of the materials. Typical examples are suitably doped alloys of PbTe and PbSe(1) or Bi_2Te_3 and Bi_2Se_3(2). Power generators having an efficiency between 8 and 10 per cent have also been made using these materials in the temperature range below 300°C. High temperature generation materials such as the mixed valence oxide compound semiconductors or systems of III–V compounds such as InAs-GaAs solid solutions show promise but have not yet been reported to have efficiencies greater than 10 per cent over temperature ranges extending above 500° C. It is important to stress, however, that materials research in this field is in its infancy. Only a small fraction of the large number of possibly useful materials has been investigated and in existing semiconductor compounds of promise such factors as oriented crystallization and defect-free crystals have been relatively neglected.

6. CONCLUSION

There seems little doubt that the impact of thermoelectric methods in refrigeration, heating, power generation, and transportation will be felt with increasing force as time proceeds. Obviously, the commercial realization of the developments in this field depends vitally on vigorous research and development in solid state materials. Moreover, research and development may be expected to be well rewarded in a field as young as this. It is thus not inappropriate to conclude this discussion by broadly summarizing the basic areas for unrelenting effort in research and engineering design, through which increasing improvement in performance can be attained.

Principal Areas for Research and Development

The development of materials of increasing figure of merit over the whole temperature range of application.

The development of techniques for further reducing semiconductor-metal contact resistances for miniaturized thermoelements.

The arrangement of mass production methods for materials fabrication and thermoelement formation.

The design of heat exchange arrangements of increased technical, weight, and cost efficiency.

REFERENCES

(1) A. F. JOFFE, "Semiconductor Thermoelements and Thermoelectric Cooling," London, Infosearch, 1957.
(2) N. FUSCHILLO, F. J. DONAHOE AND J. BIERLY, *J. Phys. Chem. Solids*, Vol. 8, p. 430 (1959).
(3) R. R. HEIKES AND C. ZENER, "Proc. Thermoelectric Conf. N.R.L. (1958)," (to be published).
(4) R. C. MILLER AND R. R. HEIKES, *J. Chem. Phys.*, Vol. 28, p. 348 (1958).

V.2
195

BOOK REVIEWS

SOME PROBLEMS OF CHEMICAL KINETICS AND REACTIVITY. VOL. I, by N. N. Semenov, translated by J. E. S. Bradley. 305 pages, diagrams, 5½ × 8½ in. New York, Pergamon Press, Inc., 1958. Price, $7.50.

Nobel prize winner Semenov has combined an earlier, limited-edition, Russian book and a collection of U.S.S.R. Academy of Science papers in preparing this edition. It contains a very critical review of the latest international literature and much information from Russian studies that were previously unpublished. This is not intended as a complete treatise of the subject. It is replete with personal views, criticisms, and suggestions. Many areas requiring research are enumerated. The author urges broader understanding of Kinetics and reactions through basic research and general investigation rather than the study of isolated observations. The book is divided into three parts.

Part I deals with radical reactions, including chapters on uniradical reactivities, alternate univalent free-radical reactions, and reactions of biradicals. Topics included in the discussion are activation and bond energies, polar and steric factors, radical decomposition and isomerization, intermediate and chain reactions, and an interesting treatment of metastable states in complex molecule photoreactions. The treatment of biradicals in the kinetics of chain reactions provides a thought-provoking excursion into this important subject.

Part II, Initiation and Termination of Chain Reactions, includes chapters on molecular dissociation and radical combination, ions of variable valency as chain initiators, and wall initiation and retardation of chain reactions. This reviewer found the last named chapter to be a highlight. The author explores many facets of wall effects including free radical generation and removal, free valencies, radical production, concepts pertaining to heterogeneous catalysis, etc. Supporting information of a quantitative nature has been combed from many sources and concisely presented.

Part III, Kinetics of Chain Reactions, treats with chain reactions competing with reactions between saturated molecules. Energy dependence, relative rates, specific decompositions and free-radical production are described.

Appendices treat with the activated complex and the quantum-mechanical calculation of activation energies. There are many references for every chapter and these are supplemented by a name and subject index.

Space limitations preclude a more detailed review. Suffice to say that scientific personnel concerned with kinetics or reaction mechanisms will find this a highly worthwhile, stimulating contribution.

ROBERT A. BAKER
The Franklin Institute Laboratories

HUMAN ENGINEERING, by Ernest J. McCormick. 467 pages, illustrations, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1958. Price, $8.00.

This text has been directed toward the consumers of human engineering recommendations—design engineers and project managers. Rather than present an oversimplified catalog of those human characteristics which are important when men act in concert with machines, the author has chosen to give the reader a certain amount of insight into the process by which these human characteristics were determined. This background material has been presented clearly and adequately. Psychological nomenclature of parochial appeal has been kept down, and recommendations are presented in a well reasoned manner.

As with any text of this nature, there are omissions and weaknesses, although the author has done a fine job of culling the unclassified literature of government reports. Nevertheless, there are strong trends in current human engineering thinking which have not been treated. One such trend is the ever increasing recognition of the complementarity of human engineering and operations research in the study of large and complex manned systems. This concept deserves a detailed

and extended coverage. By and large, however, McCormick has prepared a useful, well written book on an infant technology.

EZRA KRENDEL
The Franklin Institute Laboratories

ENGINEERING VIBRATIONS, by Lydik Jacobsen and Robert S. Ayre. 564 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1958. Price, $10.00.

This is a welcome and useful treatise on vibration analysis, particularly adaptable for directing the solution of a large class of typical engineering vibration problems.

The work is designed as a suitable course on an early graduate level and includes an extensive bibliography. The unique feature of the book is a new approach in the setting up of the differential equations and actually carrying out the solution for an important class of engineering cases. By this method, the authors have successfully developed a true physical feel of the subject. Along with the analysis, free use of graphical and tabulated methods are used, such as time displacement curves for damping ratios, and the establishment of roots in transcendental frequency equations, as well as the picturing of modes of vibrations of multi-degree-of-freedom systems.

The early chapters cover linear and nonlinear single-degree-of-freedom systems, transient and steady forced vibrations with and without damping, including approximate solutions.

Part II covers linear undamped multi-degree-of-freedom systems with various types of couplings. Numerical methods, with actual solutions for multi-mass systems, include the Holzer and the Stodola-Rayleigh methods. The shearing vibration of a six story building is discussed.

The last chapters deal with continuous systems, as the stretched string, bars with various dampings, pressure waves, and such matters as the vibration of shear beams.

The authors are well known in their contributions to vibration analysis, and in particular to vibration problems in structures, and this work reflects a senior approach to the subject. They are to be complimented in presenting a difficult subject in a clear-cut simple manner.

RUPEN EKSERGIAN
The Franklin Institute Laboratories

PRINCIPLES AND APPLICATIONS OF RANDOM NOISE THEORY, by Julius S. Bendat. 431 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1958. Price, $11.00.

The sciences of communication and control have gone through a complete evolution cycle from a primitive status, characterized by a purely deterministic approach, to the recent developments resulting from the recognition of the statistical nature of the processes involved.

One of the results of this evolution is the appearance of books intended to recapitulate and clarify the major classical original contributions illustrating at the same time the various applications.

The difficulty of locating and interpreting the numerous and not always coordinated original works makes the availability of well written books of this kind a blessing to scientists, engineers and graduate students.

Dr. Bendat's work is especially devoted to the study of the principles and applications of random processes rather than to that specific body of knowledge which goes under the name of "Information Theory". The reviewer agrees with the author's opinion that this distinction is expedient although regrettable in view of the areas of overlap between the two philosophies.

The book is divided into ten chapters and it is pleasing to find detailed summaries of the individual chapters at the beginning of the book in addition to the conventional index. Furthermore six chapters end with a concluding summary that is highly effective and clarifying. Appendices are also found at appropriately chosen points.

The first chapter gives an introduction to random processes, correlation functions and system analysis. The concepts of stationarity and ergodicity are presented in an elementary way yet with sufficient sophistication to warn the reader against indiscriminate acceptance of certain hypotheses. Systems are classified in terms of linearity, time-invariance of parameters, finiteness of operating time. A very brief treatment of superposition in the time domain and of the related frequency response functions is also given.

Chapter 2 deals with power spectra and correlation functions as time averages. The presentation of this material proceeds from the periodic to the aperiodic and, eventually,

V.2
195

to the random case. The sampling theorems are presented with special attention to their relation to the resolving power in time and frequency for any physical system. The use of correlation techniques to extract "hidden" periodicities in quasi-random processes and to obtain input-output relations for linear constant parameter systems is discussed.

Chapter 3 opens with an analysis of ensemble averages followed by a condensed exposition of probability theory culminating in the Central Limit Theorem for one random variable. The theory is extended to multivariable case with the presentation of distribution of sums, characteristic and moment generating functions. The effects of a change of variable in probability density functions is thus discussed and the chapter ends with the presentation of the multivariate normal distribution, the Central Limit Theorem in n dimensions and numerous applications to the analysis of random noise.

Chapter 4 is devoted to Optimum Linear Prediction and Filtering. Although the least-mean-square error philosophy is given due predominance a special effort is made to compare the various conditions and assumptions chosen by different authors on account of mathematical expediency or physical significance. (See Table 4-1 and Section 3 of Chapter 4). The historical approach together with frequent examples will help the reader to reach a firm understanding of the theory for any combinations of: constant, or time-varying, parameter linear systems, finite, or infinite, operating-time systems and stationary, or non-stationary, random processes.

The reader's attention is also directed to the similarities and the differences between the filter-predictor problem and that of the extraction of a signal from noise under assumptions that may differ from the least-mean-square error criterion.

Chapter 5 analyzes the exponentially damped autocorrelation function which has proven useful in many cases. Examples from thermal, tube and circuit noise, telegraph and other signals, fading phenomena in turbulent media are cited to support the above point of view. It is shown that a suitable combination of several exponential, or exponential-cosine, functions can be used to fit many physical cases. A discussion of Markoff processes closes the chapter.

Chapter 6 is devoted to analog computers designed for driving forces characterized by exponentially damped cosine functions and for random errors in the initial conditions that are assumed to be statistically independent among themselves and from the driving forces.

Chapter 7 offers quantitative criteria to estimate the statistical errors occurring in the measurement of correlation functions with finite bandwidth and finite integration time. The effect of a varying correlation interval and the concepts of "no correlation" and "false correlation" probabilities are presented.

Chapter 8 is devoted entirely to the problem of the demodulation from RF to video of signals in noise. Correlation detectors are compared to envelope ("linear") and square-law detectors. The effects of automatic gain control are briefly mentioned. The important role of a finite bandwidth–integration time product is again stressed in this Chapter.

Chapter 9 formulates the time-variable parameter filter problem in the sense of optimizing, at the filter output, an appropriate signal function or some transformation thereof when the system input is a linear combination of signals and noise. The adopted optimization criterion is to minimize the mean-square ensemble difference between the ideal and the effective outputs. This difficult problem is handled both from an analysis and a synthesis point of view and a general solution is obtained.

Chapter 10 is devoted to the so called zero-crossing problem which is important in many applications. The usefulness of the expected number of zero crossings and of the output power spectral density is considered in detail with reference to pure random-noise inputs or to combinations of signals and noise.

A bibliography of 141 references and an alphabetical index close the book.

The well coordinated presentation of known material, the original contributions, and, last but not least, the author's successful effort to organize the book in a way which should facilitate the reader's task of absorbing a large amount of knowledge, deserves the highest praise especially in an epoch like ours in which cryptic writing is often confused with virtuosity.

P. L. BARGELLINI
University of Pennsylvania

BOOK NOTES

A HISTORY OF SCIENCE, by George Sarton. 554 pages, diagrams, 6½ × 9½ in. Cambridge, Harvard University Press, 1959. Price, $11.00.

This work, completed just before Dr. Sarton's death, is his last book. Although the subtitle would indicate that the volume deals only with Greek science in the last three centuries B.C., it does in fact cover Roman culture and Latin letters, as well as the culture of Eastern Europe, Egypt, and Western Asia. The three centuries treated witnessed the establishment of geometry, astronomy, anatomy, and grammar, and the beginnings of technology and medicine. Dr. Sarton evaluates each ancient achievement in relation to what bearing it had on later achievements (in contrast to his original plan).

The book is divided into two parts. Part One treats the third century B.C., including the Alexandrian renaissance, Euclid, the beginnings of astronomy, Archimedes and the mechanics, geography, physics and technology (public works), anatomy, medicine, the library, philosophy and religion, history and language. Part Two, on the last two centuries B.C., deals with the growth of Rome and continues the study of the subjects enumerated for Part One.

This monumental work will long stand as the authoritative reference work in the field. It is a matter of deep regret that Dr. Sarton's original plan—to cover the history of science through the 19th century—cannot now be carried out.

RANDOM VIBRATION, edited by Stephen H. Crandall. v.p., diagrams, 6 × 9 in. Cambridge, The Technology Press, 1958. Price, $10.00 (paper).

This volume was prepared for use in the special summer program on Random Vibration, offered at M.I.T. in 1958. The problem of random vibration is of vital importance in the high power rocket and jet engine fields, since it severely affects not only the performance of the vehicle, but also that of the airborne electronic equipment. The several papers included in the volume describe the new concepts required to extend ordinary vibration theory into the field of random vibration. The authors also set forth the current state of the art of designing and testing equipment to withstand random vibration. The book is divided roughly into two parts: the first six chapters deal with basic concepts and background material; the last six chapters cover measurement of random vibration, design problems in aircraft and missiles subject to random vibration, excitation from jet engines, rocket motor excitation, design of simulation equipment, and mechanical loading. This book, the first of its kind, should be required reading for all designers involved in rocketry, missiles and jet engine aircraft.

COSMIC ELECTRODYNAMICS, by J. W. Dungey. 183 pages, diagrams, 5½ × 8½ in. New York, Cambridge University Press, Inc., 1958. Price, $6.00.

One of the Cambridge Monographs on Mechanics and Applied Mathematics, this small volume treats the relatively new subject of cosmic electrodynamics from a theoretical viewpoint. The author discusses the fundamental physical theories involved and explains the mathematical methods used. After a preliminary discussion, the author covers these eight general subjects: the velocity-distribution method; the motion of magnetic fields; magnetostatics; magnetodynamics; acceleration of charged particles to high energy; solar phenomena (sunspots, corona, prominences, flares *M* regions); magnetic storms and aurorae; and ionospheric electrodynamics. The monograph will be appreciated by those interested in cosmogony, astronomy, astrophysics, solar physics, atmospheric physics and others in the fields of cosmic ray work, the Earth's magnetism and particle accelerators.

V.2
195

REFLEX KLYSTRONS, by J. J. Hamilton. 260 pages, illustrations, 5½ × 8¼ in. New York, The Macmillan Co., 1959. Price, $9.00.

Written on an undergraduate level, this survey incorporates an elementary outline of the theory and operation of reflex klystrons, including the fundamentals of cavity resonators, together with a discussion of the techniques of the design, manufacture and testing of such tubes. Typical reflex klystrons, from some of the earliest production types to modern millimeter wave tubes are described and illustrated. The last two chapters cover some unconventional models and the direction in which present developments are trending.

FINE PARTICLE MEASUREMENT, by Clyde Orr, Jr. and J. M. DallaValle. 353 pages, diagrams, 5½ × 8½ in. New York, The Macmillan Co., 1959. Price, $10.50.

The authors have gathered in one volume the data on important techniques used in research laboratories for fine particle measurement. Covering theory and application, the book describes over seventy techniques currently in use for determination of size, surface and pore volume. Ten chapters deal with various phases of fine particle measurement, including microscopy and sieving, sedimentation, inertial techniques, radiation scattering and transmission, permeametry, gas adsorption, liquid-phase sorption, and X-ray evaluation of pore size.

Since fine particle measurement plays so important a role in widely diverse fields (for example, concrete, paints, foods, catalysis), the book should be of value to a wide audience. The presentation is limited to basic principles, use of apparatus, methods of data analysis, applications and limitations; it does not attempt to relate measurement with properties.

THE TECHNICAL WRITER, an aid to the presentation and production of technical literature, edited by J. W. Godfrey and G. Parr. 340 pages, illustrations, plates, 5½ × 8½ in. New York, John Wiley & Sons, Inc., 1959. Price, $8.50.

New information in all branches of science and technology needs to be disseminated as rapidly as developments take place. The designers and researchers often have not the time or perhaps the ability to write up progress in a way that could help others in industry. This creates a need for skilled technical writers who have that ability.

The term "technical literature" covers a variety of publications which have the common object of conveying information as efficiently and clearly as possible, and this is the underlying art of technical writing. In larger organizations the technical writer often forms part of a team producing technical literature. In smaller firms the engineer or designer may do his own writing and production. To give the technical writer an understanding of the technique of presentation and production is the purpose of this book. The authors have assumed that the writer is already grounded in the principles of composition, and mainly needs to avoid common pitfalls and acquire a readable style. The book also stresses that the writer's responsibility for his work carries through the entire publication process. The book familiarizes him with the processes through which his manuscript passes, aiming to promote efficient, neat and economical production.

The examples are to a large extent drawn from electrical and electronic engineering practice but the principles hold for all branches of science. The attention given here to circuit diagrams stems from the fact that in modern engineering, mechanical and electrical technologies are interdependent, and nearly all allied sciences make use of electronic circuits.

In its analysis of the forms of technical literature, the book describes the technical report, paper, article, leaflet, sales information and books. Editorial procedure, typography for technical publications, set-up of a technical publications unit, and technical authorship are each analyzed in detail.

Appendices are added covering casting-off tables, proof-correction marks, printing terms, and associations for technical writers.

THE ASTRONOMER'S UNIVERSE, by Bart J. Bok. 107 pages, illustrations, 5½ × 8½ in. New York, Cambridge University Press, 1959. Price, $3.75.

Written in non-technical terms, this text gives an account of recent developments in physics, mathematics and radio astronomy that have put new tools and methods into the hands of astronomers. Selected topics include

the measurement of size, distance and brightness of stars, problems of the solar system and the Milky Way, radioactivity and mapping the heavens. In the final chapter the author enters into the realm of theory and speculations.

ANIMAL LIFE, by Lorus J. Milne and Margery Milne. 367 pages, diagrams, illustrations, 7 × 9¾ in. New York, Prentice-Hall, Inc., 1959. Price, $6.95.

New vitality is given to the age-old science of life's origin in this richly pictorial text. It is written with great clarity in a refreshing style beginning with the basic information concerning the living processes. The importance of coordination, both chemical and nervous, in active multicellular animals is followed by contrasts of single-celled life. The dramatis personae of zoology is outlined for their economic importance.

CONSTITUTIONAL DIAGRAMS OF URANIUM AND THORIUM ALLOYS, by Frank A. Rough and Arthur A. Bauer. 153 pages, diagrams, 6 × 9 in. Reading Mass., Addison-Wesley Publishing Co. Inc., 1959. Price, $5.00.

This valuable compilation of United States and United Kingdom uranium and thorium constitutional diagrams is a compact and highly useful source of reference material for workers in the field. The book is divided into two major sections, uranium and thorium alloys, each preceded by a discussion of the transformation and melting temperatures of the base metal. The various systems are listed in alphabetical order, including both binary and ternary systems.

ELEMENTS OF PHYSICAL METALLURGY, by Albert G. Guy. 528 pages, diagrams, illustrations, 6 × 9 in. Reading, Mass., Addison-Wesley Publishing Co., Inc., 1959. Price, $9.50.

This second edition of a widely used textbook for introductory courses in physical metallurgy contains several revisions in the introduction of dislocation theory. There is additional material on electron theory of metals, in view of the increased importance of electric phenomena in semiconductors. A new chapter is included on "Phases in Metal Systems" and other material designed to bring the text completely up to date. Written especially for science and engineering or metallurgy students since it treats basic principles.

FILLER METALS FOR JOINING, by Orville T. Barnett. 244 pages, diagrams, illustrations, 6 × 9 in. New York, Reinhold Publishing Corp., 1959. Price, $7.00.

This volume summarizes the industrial practice in selecting the proper filler metal for any metal joining situation. It covers all of the ferrous and non-ferrous joining metals. Welding, brazing and soldering are treated with major emphasis on products for manual and automatic arc welding. Several chapters are devoted to arc welding electrodes and automatic welding of mild and low alloy steels and discusses materials in an area where no specifications exist as yet. The chapters on new iron powder electrodes, the latest developments in non-ferrous metals, surfacing and tungsten electrodes contain much material that has never been published before. An excellent reference for welding engineers and designers.

OCEANOGRAPHY AND MARINE BIOLOGY, by H. Barnes. 218 pages, diagrams, illustrations, 5½ × 8½ in. New York, The Macmillan Co., 1959. Price, $7.00.

In this volume the author comprehensively describes the ingenious instruments and methods by which we obtain our knowledge of the sea. Methods of taking bottom cores and of using nets, pumps, trawls, dredges and grabs are all carefully explained. A section is devoted to the properties of water itself—temperature, salinity and currents. Particular attention is given to new techniques as for instance, the bathythermograph. The last chapters deal with the most spectacular of modern developments—underwater photography and television. Aerial photography is also included as a growing contribution to oceanography. Students of marine biology will find this volume of the greatest interest while marine research workers will value the assemblance of their techniques.

V.2
195

WELDING OF PLASTICS, J. Alex Neumann. 279 pages, diagrams, illustrations, 6 × 9 in. New York, Reinhold Publishing Corp., 1959. Price, $7.25.

Various physical and chemical properties of individual plastics presently used for welded construction are completely covered in this volume, with regard to their use in chemical and allied processing. All presently known techniques of welding are covered in detail and there are many step-by-step procedures with photographs. One of the most complete corrosion resistance tables yet published is included, plus a concise guide for the choice of an appropriate plastic construction material. This text will be of substantial help to engineers and manufacturers who are interested in present and potential applications in the process industries.

LUBRICATION SCIENCE AND TECHNOLOGY, edited by John Boyd. 346 pages, diagrams, illustrations, 8¼ × 11¼ in. New York, Pergamon Press Inc., 1959. Price, $10.00.

This book represents Volume II in the Transactions of the American Society of Lubrication Engineers, constituting a valuable group of papers on widely diverse fields of lubrication. The papers are all written by leaders in their respective fields giving a source of reference not only for those concerned directly with lubrication, but also to many industrial research engineers and designers.

THE ATOM AND THE ENERGY REVOLUTION, by Norman Lansdell. 200 pages, plates, 6 × 9¼ in. New York, Philosophical Library, 1958. Price, $6.00.

In this book Lansdell presents a balanced study of the role of atomic energy in the world of today and tomorrow. He concerns himself only with the peaceful uses of energy, including the atom. The issue is the reorientation of human life and society, the impact of power development on other spheres of activity (on other industrial techniques, private industry, and the balance of world trade). The book is addressed to the specialist to acquaint him with other disciplines, to the business man who must adjust to the new industrial and commercial factors introduced by atomic energy, to the citizen to show him the social and political implications of the new sources of energy, their potentialities, demands, and possible cost.

The serious problems which must arise during the next fifty years, arising from the world pattern of energy demand, are neither the same nor equally crucial all over the world. New sources of energy are required on different scales (large stations delivering power and/or heat, medium and small fixed units, and mobile units). Such sources are wind and water power, vegetable resources, solar energy, electro-chemical generation of energy, direct conversion of radiation into electric current, nuclear fission, nuclear fusion, and the annihilation of matter. Present atomic energy developments relate mainly to large units.

The structure of the atom and its energy are analyzed. The study of the methods of releasing atomic energy is illustrated with 32 plates.

Lansdell shows the sources of natural materials for atomic energy development, the exploitation of atomic energy, the political and commercial organizations for atomic energy development, national and international, radiation risks, and insurance against them.

MATHEMATICAL METHODS OF OPERATIONS RESEARCH, by Thomas L. Saaty. 421 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc. 1959. Price, $10.00.

Operations research—the scientific approach to operational problems for the greater fulfillment of objectives—is a parallel development to an increased realization of the usefulness of staff planning and analysis functions. The implementation of these functions has led to rapid development of new methodology, and application of old methodology to new problems. The book requires for background a course in calculus, some advanced calculus and rudimentary matrix theory, and some analytical experience.

Part I of this study gives background perspective. The chapter on mathematical models illustrates, by brief discussions and examples, how the need for such tools arises, and some types of related theoretical questions.

Part II treats of optimism and some of its problems, linear and quadratic programming, and the theory of games.

Part III covers probability, applications, fundamental statistics and their problems, and a resume of the queueing theory. There are many references for further information on the subject.

Part IV is an essay covering some thoughts on creativity and elementary problems and projects. In the final analysis, the individual himself, his talents and methods of approaching problems, form the fundamental qualities required in solving problems. The analytical methods studied give a reliable approach to problems. Many of man's problems have been successfully solved by a combination of intuitive talent and a desire to accept the challenge presented by a new problem. There is a place for both the analytical and the intuitive methods of attack. The mathematical methods are given primarily to increase understanding of problems and secondarily to give solutions to these problems.

BASIC GEOLOGY FOR SCIENCE AND ENGINEERING, by Edward C. Dapples. 609 pages, illustrations, 6 × 9 in. New York, John Wiley & Sons, Inc., 1959. Price, $9.50.

This book presents a systematic analysis of the processes which are considered basic to the understanding of physical geology. It explains geologic processes, and the techniques of geologic interpretation and reconstruction of former geologic conditions.

The method of study of physical geology here is (1) observation and grouping of data, (2) classification of data into forms which can be expressed in graphs and tables to determine any measurable relationship between properties, (3) interpretation of the relationships to form significant generalizations of the behavior of geologic processes. The emphasis is on the application of geology rather than the description of localities where specific conditions are known to occur. Dapples believes that there are certain aspects of physical geology which must be considered basic to further study in the applied, descriptive or theoretical form. The pattern of his presentation is organized around empirical relationships.

The body of scientific knowledge ranges from the taxonomy of fossil organisms to theoretic analysis. The difficulty in the presentation of the subject matter of geology to the scientist and the engineer is the selection of what is to be considered the essential ingredients of the science.

Dapples studies soil materials, physical and chemical properties of rock materials, rock-forming minerals, igneous rocks, soil-forming processes, streams, shoreline processes, ground water, wind deposits, processes associated with glaciation, sedimentary rocks, crustal deformation, and metamorphism. The study is enhanced with tables for the determination of common minerals.

PLASTIC ANALYSIS OF STRUCTURES, by Philip G. Hodge, Jr. 364 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1959. Price, $10.50.

For many generations, engineers have based the analysis and design of structures on a lineal theory of elasticity. Hodge concerns himself with plastic analysis to point out some of the disadvantages of elastic-design methods. Ductile structural materials such as steel and aluminum can withstand strains much larger than those encountered within the elastic limit, and carry stresses above the yield stress (strain hardening). This ductility in redundant structures permits a redistribution of stresses beyond the elastic limit, enabling the carrying of very considerable additional loads. Although the average stress may be below the elastic limit, highly stressed local regions may become plastic, making elastic analysis frequently unrealistic.

Plastic design partially answers the objections to elasticity. It takes full advantage of ductility. Nonuniform stress distributions can be advantageously redistributed. The method is mathematically simple. Through his analysis in this book, Hodge shows that these claims are fully justified in many practical situations, and he recognizes that there are situations where plastic analysis does not fit. He explains *how* to apply plastic methods but also judges *when* these methods should be applied.

To simplify presentation, Hodge has divided the book in two parts. Part I on beams and frames treats of the application of plastic methods to frame-type structures whose principal strength lies in their resistance to bending. In addition to a strictly plastic or limit analysis, he deals with topics such as elastic plastic deformations, variable and repeated

loading, and direct procedures of design. The discussion of framed structures, where tensile stresses produce both an axial force and a bending moment, makes the transition to the combined stress problems of part II. This includes combined stresses in beams, circular plates, and circular cylindrical shells. Less complete treatment is given to more general plate and shell problems and to problems in plane stress. In conclusion, there is an introduction to some of the problems encountered in the dynamic loading of plastic structures.

As a teacher, Hodge has written this book to appeal both to the undergraduate or the first-year graduate student learning about structures for the first time, and to the practicing engineer who has completed his course work years before. With this double aim he separated the theoretical material, such as proofs and theorems, into appendices. The second part, concerned with more complex structures, is written at a more advanced level. He suggests additional references where topics have not been fully covered. Problems are offered at the end of each chapter to encourage the reader to seek out original sources.

LEGAL ASPECTS OF CONSTRUCTION, by Walter C. Sadler. 387 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1959. Price, $8.50.

This book, written for practical use in the construction industry, is focused primarily on a discussion of the rights and liabilities on construction, as a guide to help avoid legal entanglements and litigation in the areas of possible conflict of interest. The development of the book has been directed to the business of contracting, with its many entwining relationships, the relative liabilities of the many parties in the construction industry, and the property rights of abutting or adjacent parcels of land incidental to construction. The primary parties are four-fold: owners, financiers, designers, and contractors.

The business of contracting is traced through licensing of contractors, mechanic's liens, insurance, labor and management, the union contract, workmen's compensation, compensable accidents, and arbitration.

Further help in avoiding conflict of interest is given by showing court decisions in typical cases in such areas as architectural practice, architect-owner relationships, owner's responsibilities, defective design and construction, general liabilities, and professional registration.

To avoid legal entanglements "at the boundary," Sadler deals fully with boundary rights, structures at the boundary, lateral support, foundations, statutes and ordinances, money damages, and injunction and mandamus.

Sadler, being an attorney and an engineering and construction consultant, offers this book as a guide to "preventive law" for the construction industry.

PHYSICAL LAWS AND EFFECTS, by C. F. Hix, Jr. and R. P. Alley. 291 pages, diagrams, 5¾ × 9 in. New York, John Wiley & Sons, Inc., 1958. Price, $7.95.

This unusual book has been prepared to provide a centralized source of information on physical laws and effects. The authors, realizing that new ideas are frequently inspired by old laws, have gathered new laws and old ones, often getting their material from half-forgotten sources. In a brief introduction, the authors explain why such a compilation is useful. An 8-page index of physical laws and effects follows this introduction, with page numbers keyed to each, referring to the descriptions given in Chapter 3. Chapters 4 and 5 are cross references, respectively, to physical quantities and fields of science, for the material in Chapter 3. A final chapter contains several rules, effects and laws which, because the authors do not vouch for their accuracy, are not included in the main body of the text.

PUBLICATIONS RECEIVED

ELECTRONIC FUNDAMENTALS AND APPLICATIONS, by John D. Ryder. Second edition, 721 pages, diagrams, 6 × 9 in. New York, Prentice-Hall, Inc., 1959. Price, $13.35.

PLANE TRIGONOMETRY, by Gordon Fuller. Second edition, 281 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1959. Price, $4.75.

FUNDAMENTALS OF ENGINEERING DRAWING, by Warren J. Luzadder. Fourth edition, thoroughly revised, 720 pages, diagrams, 6 × 9 in, New York, Prentice-Hall, Inc., 1959. Price, $10.00.

GENERAL CHEMISTRY, by William H. Nebergall and Frederic C. Schmidt. Alternate edition of "General Chemistry" by the same authors, 723 pages, illustrations, 7 × 9 in. Boston, D. C. Heath and Company, 1959. Price, $7.25.

ELECTRICAL ENGINEERING: THEORY AND PRACTICE, by William H. Erickson and Nelson H. Bryant. Second edition, 614 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1959 Price, $8.00.

MECHANICS. PART II: DYNAMICS, by J. L. Meriam. Second edition, 420 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1959. Price, $5.00.

SEMICONDUCTOR ABSTRACTS, compiled by Battelle Memorial Institute. Volume IV—1956 issue, edited by E. Paskell. 456 pages, 8½ × 11 in. New York, John Wiley & Sons, Inc., 1959. Price, $12.00.

HI-FI FURNITURE, by Jeff Markell. 224 pages, illustrations, 5½ × 8½ in. New York, Gernsback Library, Inc., 1959. Price, $2.90 (paper); $5.00 (hard cover).

INTRODUCTION TO SYMBOLIC LOGIC AND ITS APPLICATIONS, by Rudolf Carnap. Translation of "Einführung in die symbolische Logik," revised and published for the first time in English. 241 pages, 5½ × 8 in. New York, Dover Publications, Inc., 1958. Price, $1.85 (paper).

TIME'S ARROW AND EVOLUTION, by Harold F. Blum. Second edition, 224 pages, 6 × 9 in. Princeton, Princeton University Press, 1959. Price, $1.75 (paper).

ORGANIC SYNTHESES. VOLUME 38. 1958. 120 pages, diagrams, 5¾ × 9 in. New York, John Wiley & Sons, Inc., 1958. Price, $4.00.

V.2
195

CURRENT TOPICS

Durable Pleating and Creasing Process for Wool.—Wool trousers and skirts with creases and pleats that never lose their press will soon be available all over America. Development work on volume production techniques for applying the new Si-Ro-Set process for the durable pleating and creasing of woolens and worsteds has now been completed, according to Max F. Schmitt, president of The Wool Bureau, Inc.

Experimental work on the adaptation of the process to U. S. factory production has been going on in the United States since early 1958, shortly after its development in Australia. Si-Ro-Set was the result of research conducted by the Commonwealth Scientific and Industrial Research Organization in that country.

Mr. Schmitt said that two types of test production have been carried on under the supervision of Dr. Gerald Laxer, The Wool Bureau's research director.

One method involves the manual operation of conventional spraying and pressing equipment specially adapted for use with the new process. The other method utilizes new, automatic spraying and pressing equipment developed specially for the Si-Ro-Set process under Dr. Laxer's direction.

Full information on how to use Si-Ro-Set was disclosed in March, in a technical booklet prepared for release to the industry. The Wool Bureau has also produced a motion picture to serve as a graphic method of instruction in the application of the process.

Personal instruction for clothing manufacturers is available through Dr. Laxer and Charles Macaluso at The Wool Bureau. Mr. Macaluso was added to the technical staff of The Wool Bureau at the end of 1958 to aid in the systematic introduction of Si-Ro-Set to manufacturers of woolen and worsted garments in the United States.

Si-Ro-Set involves spraying garments with a certified Si-Ro-Set solution which is similar to one used in the "home permanent" wave. It imparts durable pleats or creases to wool apparel without altering the hand or life of the fabrics. Garments treated with the process will hold the crease through rain, high humidity, dry cleaning or total immersion.

Mr. Schmitt said that chemical manufacturers signing agreements with the Commonwealth Scientific and Industrial Research Organization for use of the certification mark Si-Ro-Set may do so without charge. Garment manufacturers purchasing Si-Ro-Set concentrate will be automatically given a label license to use the concentrate in accordance with the process or processes covered by any patents or patent applications owned by the CSIRO, without charge.

Unique Shelter for Use with Jupiter Missile.—A shelter that opens like the petals of a flower has been developed by the U. S. Army Engineer Research and Development Laboratories, Fort Belvoir, Virginia, for use with the Jupiter intermediate-range ballistic missile. The weatherproof shelter is designed to cover the lower portion of the rocket, creating a watertight seal that protects instruments

and equipment until the rocket is ready to be fired.

The shelter has twelve 30-ft. aluminum panels which are shaped like the petals of a flower. After installation, they radiate from a circle around the missile. When preliminary adjustments on the missile are completed, synchronized motors raise the panels, closing them around the rocket. When the missile is ready for firing, electronic equipment takes over, causing the panels to open up and return to the ground.

Doors in the 30-ft. sections allow workers to go inside for inspection of the launching apparatus. Other openings carry fuel lines to the missile. The interior of the shelter is coated with a paint capable of resisting temperatures of over 1000 F.

The engineering firm of Barnes and Reinecke, Incorporated, Chicago, Illinois, holds the principal Army contract for production of the shelter.

Auto-Optic Recording Pyrometer.—An optical recording pyrometer that can measure extremely high temperatures with milliseconds response time has been announced by Avco Research and Advanced Development Division of Wilmington, Mass. Developed to measure the high surface temperatures of sample materials exposed to Avco's 10,000 C. electric-arc air-heater, the new pyrometer may be operated unattended when conditions are hazardous to personnel.

Even under adverse conditions such as an electric-arc experiment, the pyrometer has given reliable measurements of sample brightness temperatures within ±20 C. It accurately measures surface changes in temperature over an area of $\frac{3}{8}$-in. diameter in the range of 1400 C. to 3000 C. A recorder may be plugged into the pyrometer to provide a complete record of temperature versus time for the entire operation.

Called the Auto-Optic Recording Pyrometer, the radiation-sensitive element is a high-vacuum phototube. Monochromatic light of wavelength 0.65 micron is focused upon an opaque plate having a small opening which admits the radiant energy to the phototube. The phototube output is amplified by a highly stable amplifier mounted in the pyrometer housing. The output may be read on the built-in meter or from an external recorder.

A remote recorder attached to the pyrometer is used by Avco to study high-temperature properties of new materials being developed for possible use on space vehicles.

The instrument is calibrated by comparing the output with the known brightness temperature of a standard tungsten lamp.

TASI Will Double the Capacity of Submarine Telephone Cables.—A new system which will increase the number of telephone conversations carried by underseas telephone cables perhaps by as much as two times is under development at Bell Telephone Laboratories. The system is called TASI, which stands for Time Assignment Speech Interpolation.

In a normal two-way telephone conversation, one person on the average will be speaking only half the time, and even while he is speaking, there will be significant gaps and pauses in his speech. Thus, if the two directions of transmission are separated, each transmission path is idle, on the average, more than half the time. If two conversations are interlaced to take advantage of these gaps, greater use can be made of existing transmission facilities. It is this interlacing process which leads to the designation Time Assignment Speech Interpolation.

In practice, the system would not work with only two talkers on one line since they would frequently be speaking at the same time. However, where a larger number of channels is available, such as a submarine telephone cable, an averaging effect occurs so that at any instant there is a greater probability of sufficient "free time" being available to accommodate the larger number of conversations. Increasing the capacity of a 36-channel system to 72 channels is more feasible with TASI than doubling the capacity of a 5-channel system.

TASI is essentially a group of high-speed switches. If, for example, 36 cable channels were available, 72 talkers could be connected. When there are more talkers than channels, the equipment will connect talkers who become active by disconnecting talkers who are silent at that moment. In turn, this disconnected talker will be assigned another momentarily inactive channel when he starts to speak again. A talker will be disconnected only when he is silent.

When a talker starts talking, his voice actuates a speech detector. The speech detectors are scanned by a control circuit similar to a modern digital computer. When a talker becomes "active" this control circuit initiates a coded tone burst consisting of a group of four audio tones which precedes the voice over an available cable channel. After the tone burst, which lasts only 10 milliseconds, the control circuit connects the talker to the same channel. The coded tones operate switches to connect the talker to the proper line at the receiving end. The tones are not heard since the listener is not connected while they are being transmitted. When a talker is not "active" and his channel is needed, another coded tone burst is transmitted over a separate signaling channel and severs the connection.

Switching of talkspurts from one channel to another is accomplished in a few milliseconds by a time-division switch. Speech is sampled for about two microseconds and the resulting pulses steered to the appropriate idle channel during the sampling by the selective operation of transistor "gates" in each channel. This short interval of sampling makes it possible to sample all active talkers 8000 times a second and then reconstruct the speech from the samples before it is transmitted over the undersea cable.

The signaling system in TASI keeps the receiving end informed of the connections that the transmitting end has established. Four groups of audio tones are employed for signaling purposes—four tones in each of three groups and three in the fourth group. Each signal comprises one and only one tone from each group—if more or less are present, an error is indicated and appropriate steps taken to correct it. "Connect" signals which precede the voice signal at the beginning of the speaker's talkspurt are sent over the same channel as the speech. Disconnect and connection checking signals are sent over a separate channel used only for this purpose.

All of the circuits for TASI are completely transistorized. Terminals for doubling the capacity of the Transatlantic cable will require several thousand transistors of four different types, and tens of thousands of semiconductor diodes and passive components.

Microscope Color TV System.—Specialized medical television equipment, including a highly sensitive camera tube that can see in the dark and a compact transistorized color TV pickup system developed by the Radio Corporation of America, was demon-

strated recently to leaders in medical education, at a two-day conference on television uses in post-graduate medical education. The sessions were held at the Clinical Center of the National Institutes of Health, under joint sponsorship of the American Academy of General Practice and the Institute for the Advancement of Medical Communication, Inc.

The demonstration program was arranged by Dr. Murray Brown, Chief of Clinical Professional Education at the National Institutes of Health, who has directed experiments at the Institutes with various television techniques.

Demonstrations included a showing of results obtained by scientists with television pickup at very low light levels by color and black-and-white systems employing an RCA-developed wide-space Image Orthicon camera tube. The tube is so sensitive it can view scenes in surroundings which appear entirely dark to the human eye.

One example shown was a color film recording of the retina of the human eye, achieved with a system employing the wide-space Image Orthicon to permit a light level low enough to avoid disturbing the patient. In another application, the extremely sensitive tube was employed to view a demonstration of intricate apparatus, illuminated only by an ordinary 75-watt light bulb, showing application of such pickup equipment in normal surroundings without the high-level lighting ordinarily required for studio-type television pickup systems. The subject was an animal specimen within a germ-free tank, and the televised images were displayed by a projection system onto a large screen adjacent to the tank, making them clearly visible to the audience.

The conferees also were shown a fully transistorized miniature color television pickup system developed at the RCA Laboratories by a research group under Dr. V. K. Zworykin, Honorary Vice President of RCA. The 20-lb. camera, employing developmental ½-in. Vidicon pickup tubes, picked up color images of subjects under a light microscope for display on standard color television receivers. The experimental system, contained in two compact units weighing only 65 lb is designed for closed-circuit use in industry, education, defense and research.

Dr. Zworykin pointed out that the system includes equipment enabling it to produce a standard color television broadcast signal permitting its possible use in picking up scenes for transmission over the air, as well as over closed-circuit systems. He added that the equipment was to be used in April by Smith, Kline and French Laboratories, Philadelphia, to transmit color microscope images to five medical schools in the Philadelphia area during a closed-circuit television lecture series.

Stereo Sound Broadcasting Now Compatible with Single Channel Reception.—Several commercial radio and television stations have recently begun offering stereophonic sound programs, by broadcasting over two separate sound channels. In various experimental arrangements, the two channels required have been selected from different combinations of AM, FM, and TV channels. The listener then spaces the appropriate receivers properly in his home. Results have been sufficiently favorable that more and more broadcasters are now considering offering stereo programming.

The major obstacle to vastly increased use of this type of stereo broadcasting, however, is the person

V.2
195

who listens with only one receiver. If the broadcaster tries for the full stereo effect, the sound the single channel listener hears comes from only one of the two widely-spaced microphone pickups, and he misses a portion of the program. What he does receive is poorly balanced, because of the placement of the pickups in relation to the sound sources. The broadcaster has had to dilute the stereo effect in order to preserve satisfactory reception for the single channel listener.

Now this single channel problem may be eliminated without affecting the stereo listener, through the use of a new "compatibility" circuit which has been developed by F. K. Becker of Bell Telephone Laboratories. The circuit depends for success on a psycho-acoustic phenomenon known as the "Precedence Effect." This effect operates in such a manner that when a single sound is reproduced through two separate loudspeakers, but is delayed several milliseconds (thousandths of a second) in one, the listener will "hear" the sound as if it came only from the speaker from which he heard it first. He will judge the second loudspeaker to be silent. The amount of the Precedence Effect depends somewhat on the length of delay built-in; at 10 milliseconds, the sense of direction for an average observer is as if the echo were 8 to 10 decibels softer than the sound preceding it.

In the new development, the circuits between the microphone pickups and their corresponding radio or TV transmitters are cross connected through two delay lines, each with its own buffer amplifier. Because of these cross connections, music or voice signals from the left microphone are transmitted directly to the left loudspeaker in the listener's home, while the same signal is slightly delayed before reaching the speaker to his right. The stereo listener will hear the sound as if it came only from the left loudspeaker because of the Precedence Effect. Conversely, the sound from the right microphone goes direct to the right speaker, but is delayed before reaching the left speaker, and is therefore unheard. Thus, the brain of the stereo listener localizes the sound he hears as coming direct from each of his two speakers, and full stereophonic effect is maintained.

However, monophonic reception is completely compatible with this, since a listener to each single channel hears the total sound from both microphones in a balanced reproduction. The slight delay built in apparently does not affect his reception at all, according to subjective tests performed at Bell Laboratories.

Typical operating conditions in listening tests conducted today had a time delay of 10 milliseconds, with the volume of the delayed channel 1½ db softer than that of the direct. (The Precedence Effect, which was first discovered in 1933, operates over a delay range of from 5 to 35 milliseconds.)

This development should make it possible for many more broadcasters to offer double-channel stereo programming without diluting the stereo effect or penalizing the single channel listener, who will make up the majority of their audience.

Solar Cookers for Underprivileged. —University of Wisconsin anthropologists are stationed in the Arizona desert and in central Mexico to watch Indians bake their tortillas on solar cookers developed by the engineers in the UW Solar Energy Laboratory.

They hope to find out for the engineers if the cooker, carefully built and

tested in the laboratory, is effective in constant daily use; and if it is, whether the people for whom it was developed, who live in areas where only sunlight and poverty are plentiful, will be willing to change their customs and use it every day.

Four Indian tribes in Arizona have been cooperating with the anthropologists in the tests since last summer, and during the second semester two selected villages in Mexico will also be under constant observation.

Although scores of the cookers have been distributed in various underdeveloped areas of the world, no attempts have been made until now at extensive day-by-day studies of their efficiency and social acceptability, according to Prof. Milton Barnett, who heads the project for the anthropologists.

If the tests are successful, both scholars and Indians will profit: the anthropologists in valuable material on the hows and whys of cultural change, the engineers in the satisfaction of building something of benefit to under-privileged mankind, and the Indians—Mohaves, Chemehuevi, Hopi, and Navajo in Arizona, and the Otomi in Mexico—with a cheap, quick method of cooking their food.

The anthropologists spend a preliminary period with the Indians to gather basic ethnographic data on their physical characteristics and social and cultural traits and to establish friendly relations with the families who will participate in the testing. Only then do they bring in the saucer-shaped and plastic-coated cookers which collect and concentrate the rays of the sun on a pot or pan mounted on an attached rack.

The heat is intense enough to boil a quart of water in 10 minutes. They carefully demonstrate the cookers until the natives have mastered the operation, and then they observe personal and cultural reactions to them.

"The men in Engineering are on their toes," Barnett says. "Because one of the Hopi women suggested that hot water for laundry was a problem, they designed a solar unit to heat water. They are working on a design for a do-it-yourself kit to cut down the cost of the heater and make it more available to people short of cash; and two students in chemical engineering are designing a field sterilizer for doctors and nurses to use in isolated places.

The UW Solar Energy Laboratory is one of the few in the world carrying on a broad research program in use of sunlight as a source of energy. The program was started in 1955 under direction of Prof. Farrington Daniels, chairman of the chemistry department and internationally recognized authority on solar energy.

New Error-Correction Code for Bursts of Errors.—Lightning flashes and other electrical disturbances which cause static and noise on communication lines may result in groups or bursts of errors in the data being transmitted over these lines. These errors can be largely eliminated by a new error-correcting code developed by Dr. D. W. Hagelbarger of Bell Telephone Laboratories. The terminal equipment required for this new code is simple and inexpensive, and synchronization is relatively easy to maintain.

Previous error detecting and correcting codes either could not handle adjacent errors, or required complicated terminal equipment and presented difficult synchronization problems. The new code is applicable to systems where the data must be accepted and delivered continuously, rather than in batches, according to the inventor.

V.2
195

Error correcting systems are not required in the transmission of speech or television signals, because the information is basically analog in character, and high degrees of redundancy exist. This reduces the detrimental effect of any errors which may occur. However, in telegraph lines, or in data transmission such as the Dataphone system recently announced by Bell Laboratories, information is transmitted in a discrete pulse-no pulse form, and mutilation of a single digit could throw a complete set of data in doubt. In this type of transmission, bursts of errors could be critical, and correction methods become an essential part of transmission equipment.

Terminal equipment using the new code can be designed to handle practically any length error burst which system analysis indicates is required. In general, the shorter the maximum burst length to be corrected, the smaller and simpler will be the terminal equipment. Also, a short burst length will result in less guard space, or "clean data section," which must follow the burst before another group of errors can be corrected.

In its simplest form, the coding system uses alternate data digits and check digits, giving a redundancy of one-half. If such a system is designed to correct error bursts of length six or less (three data digits and three check digits), the encoder consists basically of a shift register of length seven. The data digits enter the first position, and are shifted through the register before being transmitted. At each shift, a check digit is computed which makes parity (the sum of 1's) of the check digit and the data digits in the 1st and 4th positions even (zero or two). This check digit is transmitted soon after it is computed, preceding the transmission of its nearest associated message digit by seven digits. Data and check digits emerge from the coding system alternately, forming the coded message.

At the decoder, the received message is separated into check and data digits, each group entering a separate shift register. There are two copies of the parity circuits: one checks the parity relationship among data digit 1, data digit 4, and checking digit 7, while the second checks the parity among data digits 4 and 7 and check digit 10. The decoding rule is: whenever *both* parity circuits fail (by indicating an uneven parity), the data digit in position 4 is changed while being shifted to position 5. If only one parity check fails, no change is made.

With the redundancy of $\frac{1}{2}$, the synchronization problem is minimized, since the data digits and check digits alternate.

After any burst of length 6 or less, a 19-digit errorless message or guard space is enough to fill the decoder shift registers completely and prepare the decoder for another burst.

The redundancy of these codes can be reduced if desired, or they can be used to correct for bursts longer than 6, without difficulty, but with simultaneous increases in terminal equipment costs and in the guard space required.

The decoder can be equipped to *detect* bursts longer than those it is designed to *correct*. For instance, a code to correct error bursts of length six will detect all but seven of the 512 possible bursts of length 9 or less, and operate a warning device.

Evaluation of the new code is now being conducted under transmission line conditions by Dr. F. E. Froehlich of Bell Laboratories. Initial results indicate substantial improvement factors in code redundancies of $\frac{1}{2}$ and $\frac{1}{4}$. The evaluation program is continuing.

The International Yard and Pound.—The Directors of the following standards laboratories—Applied Physics Division, National Research Council, Ottawa (Canada), Dominion Physical Laboratory, Lower Hutt (New Zealand), National Bureau of Standards, Washington (United States of America), National Physical Laboratory, Teddington (United Kingdom), National Physical Research Laboratory, Pretoria (South Africa), National Standards Laboratory, Sydney (Australia)—have discussed the existing differences between the values assigned to the yard and to the pound in different countries. To secure identical values for each of these units in precise measurements for science and technology, it has been agreed to adopt an international yard and an international pound having the following definition: the international yard equals 0.9144 meter; the international pound equals 0.453 592 37 kilogram.

It has also been agreed that, unless otherwise required, all nonmetric calibrations carried out by the above laboratories for science and technology on and after July 1, 1959, will be made in terms of the international units as defined above or their multiples or submultiples.

It will be noted that the International Inch is approximately 2 parts per million shorter than the inch presently used by the National Bureau of Standards, and somewhat less than 2 parts per million longer than the inch now used by the National Physical Laboratory. To avoid possible confusion, during the transition period National Bureau of Standards calibrations of length or mass expressed in English units will embody a statement indicating clearly the unit which has been used if the choice introduces a significant difference in the calibration values. Furthermore, if the accuracy of the calibration is such that the certified values would be the same in either "International" units or the older units, the qualifying adjective "International" will not be used, that is, the values will be expressed, for example, as so many inches or pounds.

This unit shall be referred to as the American Survey Foot. Inasmuch as there is little or no interchange of survey data, where the foot measurements are used, with industrial and scientific data, where the international units will be used, it is anticipated that no confusion will result from this dual usage. For example, base line surveys which might enter into a velocity of light determination would invariably be made in terms of meters.

The values of the pounds currently in use in the United States, United Kingdom and Canada are as follows: 1 U. S. pound = 0.453 592 4277 kg; 1 British pound = 0.453 592 338 kg; 1 Canadian pound = 0.453 592 43 kg; 1 International pound = 0.453 592 37 kg.

The relative differences in the various pounds are substantially less than in the yards but since masses can be measured with greater accuracy than lengths, the differences can be significant. The present British pound is about one part in ten million smaller than the international pound, whereas the U. S. and Canadian pounds are about one and one-half parts in ten million larger.

The conversion factor for the international pound was selected so as to be exactly divisible by 7 to give the following value for the grain: 1 International grain = 0.06479891 gram. The grain is the common unit in avoirdupois, apothecary and troy pounds. There are 7,000 grains in the avoirdupois pound, and 5,760 grains in both the apothecary pound and the troy pound.

V.2
195